UNIVERSITE JOSEPH FOURIER GRENOBLE

L'EVALUATION FONCTIONNELLE DE LA MAIN : REVUE DE LA LITTERATURE, ETAT DES LIEUX ET PERSPECTIVES.

MEMOIRE POUR LE DIPLÔME INTERUNIVERSITAIRE DE REEDUCATION ET APPAREILLAGE DE LA MAIN
PRESENTE LE 25 JUIN 2005
PAR
Le Dr Michel KONZELMANN Né le 13 Août 1963 au Raincy

JURY : Pr MOUTET (Grenoble)
Pr HERZBERG (Lyon)
Md Blanchard (Grenoble)

REMERCIEMENTS :

- **Au Dr Gilles RIVIER**
  Chef du service de réadaptation générale de la clinique romande de réadaptation SUVA care, Sion, Suisse
  Pour m'avoir donné l'idée un peu folle de ce mémoire.
  Pour la relecture du chapitre V.

- **Au Pr François MOUTET**
  Chef du service de chirurgie plastique au CHU de Grenoble
  Pour avoir accepté le sujet de ce mémoire et pour son enseignement lors du DIU sur l'appareillage et la rééducation de la main.

- **Au Pr Olivier DERIAZ**
  Chef du service de recherche et qualité de la clinique romande de réadaptation SUVA care, Sion, Suisse.
  Pour la relecture de ce mémoire et ses conseils sur le chapitre consacré aux statistiques.

- **A Mr Pierre-Alain COMTE**
  Chef du service d'ergothérapie de la clinique romande de réadaptation SUVA care, Sion, Suisse
  Pour avoir relu patiemment et critiqué de façon constructive ce mémoire.

- **A Mr Alain QUIRIGHETTI**
  Physiothérapeute cadre à la clinique romande de réadaptation SUVA care, Sion, Suisse
  Pour m'avoir fourni la matière du chapitre sur l'ECF.

PLAN DU MEMOIRE

I INTRODUCTION

Qu'est ce que l'évaluation fonctionnelle de la main ?
La réponse à cette question n'est pas univoque et il n'existe pas de consensus sur ce que comprend une évaluation fonctionnelle (9, 43,51). Les fonctions de la main et les procédures d'évaluation sont multiples et le terme d'évaluation fonctionnelle est utilisé fréquemment de manière inconséquente (51).
Bear-Lehman (9) a défini l'évaluation de la main comme une technique d'examen grâce à laquelle la fonction manuelle du patient est quantifiée et jugée.
Xenard (96), pense que le concept d'évaluation fonctionnelle de la main pourrait être remplacé par le concept de dextérité. Cette dextérité est globalement l'habilité à utiliser ses mains et elle dépend des préhensions, de la coordination motrice et des apprentissages à la manipulation d'objets. Un chapitre concernera les différentes dextérités.
Kimmerle et collaborateurs (51), ont développé en 2003 un modèle intitulé le « répertoire fonctionnel de la main ». Ce modèle comporte quatre composants clés : les contraintes personnelles (physiques et psychologiques), les rôles tenus par la main (uni ou bi manuels), les actions de la main (extension de la main, serrage et manipulation d'objets) et les paramètres des tâches que la main doit effectuer. Ce concept est intéressant et regroupe l'ensemble des paramètres qu'une évaluation moderne de la main devrait comporter. Je reviendrai plus loin sur ce modèle. J'essaierai de définir ce qu'est une bonne évaluation fonctionnelle de la main et ce qu'elle devrait comporter.
La deuxième question posée est quoi évaluer et comment le faire ?
De façon classique, l'évaluation fonctionnelle de la main s'intéressait surtout aux données physiques : mobilités articulaires, force de préhension, évaluation musculaire et tendineuse, troubles de la sensibilité, œdème, (5, 9, 30, 49, 67, 68,78, 96). On sait que certains de ces tests manquent de reproductibilité et de sensibilité et qu'il n'existe pas de corrélation entre les altérations physiques observées et les capacités fonctionnelles du patient évalué (9).
Depuis 25 ans, de nombreux outils dédiés à la main et au poignet se sont développés : étude de la dextérité, des activités de la vie quotidienne (AVQ) et questionnaires d'autoévaluation. Ces outils dont beaucoup sont validés et reproductibles font partie maintenant de l'arsenal des moyens mis à disposition du thérapeute pour une bonne évaluation. Ils sont considérés comme étant plus objectifs que les tests classiques. De tels outils sont indispensables pour apprécier l'issue (le résultat = outcome en anglais) d'une thérapie, pour pouvoir comparer l'efficacité de différentes thérapies entre elles et d'une façon générale pour parler un langage commun aux différents thérapeutes de différents pays. Bien sûr, l'évaluation classique reste irremplaçable et ces nouveaux outils doivent en être le complément et aider les thérapeutes à préciser les capacités et les limitations fonctionnelles des patients.
Cette approche évaluative centrée sur le patient débouche sur une prise en charge thérapeutique personnalisée ayant pour but de retrouver un niveau optimal de fonctionnement dans les tâches de la vie quotidienne (9).
Le but de ce travail est de faire une revue de la littérature et un état actuel de l'évaluation fonctionnelle de la main et du poignet. J'ai volontairement laissé de côté l'évaluation physique (mobilité, force, sensibilité) pour m'intéresser aux tests de dextérité et de la vie quotidienne et aux questionnaires d'autoévaluation. En effet, c'est dans ce domaine que depuis 25 ans les travaux les plus intéressants sont publiés. Mon étude de la littérature s'est intéressée à la période 1960-2005 et j'ai essayé de faire une synthèse de ces différents travaux en me focalisant sur ceux dont les qualités sont reconnues et validées.

Dans une première partie je ferai un rappel sur les qualités statistiques que doivent avoir les différents tests et questionnaires pour pouvoir être valides, fiables et utilisables en pratique courante. Ceci permettra de bien comprendre les autres parties.
Dans la deuxième partie, j'aborderai les principaux tests classiques de dextérité, et j'essayerai de dégager une attitude pratique quant à leur intérêt et leur utilisation.
La troisième partie sera consacrée aux tests de dextérité de la vie quotidienne et nous verrons qu'ils sont finalement peu nombreux à être utilisables.
La quatrième partie s'intéressera aux questionnaires d'autoévaluation qui sont nombreux, faciles à utiliser et à mon avis indispensables dans une évaluation moderne.
Dans une dernière partie, j'aborderai l'évaluation fonctionnelle de la main par le versant professionnel en rapportant notre expérience pratique à la clinique romande de réadaptation.
Enfin, en conclusion, j'essayerai de définir l'évaluation fonctionnelle de la main telle qu'elle devrait être à mon avis conçue en 2005, pour des lésions de l'appareil locomoteur.

II. NOTIONS DE STATISTIQUES.

Pour comprendre des études parfois complexes, il me semble important de préciser quelques notions statistiques indispensables, qui sont valables aussi bien pour les tests de dextérité ou de vie quotidienne que pour les questionnaires d'autoévaluation.
Fess (30) dans un article de 1986 fait un plaidoyer pour que les instruments de mesure d'évaluation de la main soit valides et fiables, ce qui à l'époque n'était le cas que pour un petit nombre d'entre eux.
Luquet et collaborateurs en 1996 (58), ont précisé le concept d'unidimensionnalité d'une mesure fonctionnelle du membre supérieur.
Rudman et Hannah (77), ont décrit dans un excellent article de 1998, quelles sont les qualités qu'un instrument d'évaluation de la main doit avoir.
Enfin, Bialocerkowski et collaborateurs en 2000 (14), ont fait une revue du contenu et de la qualité des instruments de mesure du poignet.
De ces articles, il ressort que cinq grandes catégories de propriétés permettent de se faire une idée de la qualité d'un instrument d'évaluation : l'utilité clinique, la standardisation, le but de l'instrument, les qualités psychométriques (fiabilité et validité) et le point de vue du patient.
L'utilisateur de ces instruments doit clairement définir ce qui doit être mesuré et pourquoi, pour sélectionner le meilleur instrument d'évaluation.
Nous allons étudier en détails chacune des cinq catégories.

II.1 Utilité clinique.

Cette catégorie comprend quatre sous-catégories.

II.1.1 Applicabilité clinique

Elle est définie comme l'utilité de l'information fournie par l'instrument pour les buts qu'il s'est fixé. Le type des résultats, le type des tâches, la méthode d'administration et l'interprétation des résultats sont indicatifs de l'utilité de l'instrument.
Le type des résultats peut être **quantitatif** ce qui réduit l'erreur subjective et fournit une méthode standardisée, ou **qualitatif** ce qui permet d'aller plus loin dans l'évaluation de la complexité et de la qualité de la fonction de la main.
Le type des tâches : l'instrument se doit d'inclure des tâches de la vie quotidienne pour permettre une interprétation de la capacité du patient à fonctionner dans la vie quotidienne. En particulier une évaluation de la fonction bi manuelle est importante dans un tel instrument.
La méthode d'administration doit être adaptée à ce qu'on veut mesurer et au type de patients évalués (autoévaluation, interview, observation directe).
L'interprétation des résultats doit être en accord avec la philosophie du test ou du questionnaire. Dans le cas précis d'une évaluation fonctionnelle de la main, l'interprétation devrait faciliter une analyse extensive de la fonction de la main.

II.1.2 Spécificité.

La spécificité d'un instrument se réfère à la population utilisée lors du développement et de l'application de l'instrument en question. Un instrument développé pour un groupe donné ne sera pas forcément applicable à un autre groupe. En d'autres termes, la validité et la fiabilité d'un instrument avec une population donnée n'implique pas qu'il soit fiable et valide pour une autre population.

II.1.3 Disponibilité et temps consommé.

Ce sont des considérations pratiques. La disponibilité d'un instrument peut être limitée par le coût, la langue de publication, le fait qu'il soit préfabriqué ou bien dans le domaine public. Le temps consommé inclus le temps d'entraînement pour le thérapeute, le temps d'administration et de calcul du score et le temps d'interprétation des résultats.

II.1.4 Acceptabilité par le patient.

Cette acceptabilité peut influencer la façon et la bonne volonté du patient de participer à l'évaluation. Il faut que l'instrument soit facilement compris par le patient.

II.2 Standardisation.

Un instrument standardisé doit avoir des instructions précises pour l'administration, le calcul du score et l'interprétation des résultats, des critères pour le matériel utilisé et une validité et fiabilité documentées. Les instructions pour le test doivent être faciles à comprendre. Toutes ces informations devraient être publiées dans un manuel d'utilisation.

II.3 Buts- Intentions.

Il existe trois buts pour l'évaluation : décrire un état, prédire une issue ou un résultat et évaluer le changement.
Trois types d'instruments correspondent à ces différents buts : descriptif, prédictif et évaluatif.
Un instrument descriptif décrit l'état d'un patient à un moment donné et utilise souvent des comparaisons avec des normes.
Un instrument prédictif classifie des individus et est utilisé pour prédire un résultat spécifique.
Un instrument évaluatif est fabriqué pour mesurer le changement d'un individu ou d'un groupe au cours du temps.
Le but d'un instrument doit être établi clairement dans un manuel ou des articles qui font partie du développement de l'instrument. Il est très important de clarifier le but de l'instrument, afin de pouvoir l'examiner de manière critique avec des critères appropriés. Cela permet aussi de s'assurer que l'instrument a bien été développé pour un but précis dans une situation clinique spécifique.

II.4 Propriétés psychométriques.

Elles comportent quatre éléments : items, fiabilité, validité et normes. Il est essentiel qu'un instrument de mesure soit fiable et validé. Si tel n'est pas le cas, le clinicien qui utilise l'instrument de mesure n'a pas la certitude que ce qu'il mesure est juste et surtout que ce qu'il mesure est bien ce qu'il avait envie de mesurer.
L'évaluation critique des propriétés psychométriques implique l'étude du but du test, les résultats statistiques et la qualité de la recherche. Les propriétés psychométriques d'un instrument doivent être réévaluées chaque fois que cet instrument est utilisé dans un nouveau contexte ou avec un groupe de patient différent.

II.4.1 Constructions des items.

Les items de l'instrument doivent correspondre aux buts recherchés et les auteurs doivent préciser le rationnel utilisé pour la sélection des items. Une évaluation descriptive doit inclure un large choix d'items pour différencier les personnes avec ou sans lésions de la main.
Une évaluation « évaluative » doit inclure seulement des items sensibles au changement dans la fonction de la main.

II.4.2 Fiabilité (reliability en anglais).

La fiabilité d'un instrument est l'étendue avec laquelle cet instrument produit des résultats cohérents et reproductibles. Un instrument fiable a des petites erreurs de mesures et est stable à travers les évaluateurs, les sujets évalués et le temps.
Il existe quatre types de fiabilité : interindividuelle, intra individuelle, temporelle ou test-retest et cohérence (consistency) interne.
Pour l'instrument descriptif, la cohérence interne est la plus importante. Pour les instruments évaluatif et prédictif, les fiabilités intra, inter individuelles et test-retest sont fondamentales.
Les résultats statistiques de fiabilité sont exprimés en termes de coefficients de corrélations : *coefficient de corrélation intra classe, coefficient de corrélation de PEARSON et Cohen kappa*. Les valeurs de ces coefficients varient de 0 à 1 avec une fiabilité maximale si le coefficient est proche de 1. La signification de ce coefficient de corrélation dépend du type de résultat et de la taille de l'échantillon.

Fiabilité Inter observateurs : C'est l'étendue avec laquelle un instrument produit des scores cohérents lorsqu'il est utilisé par différents évaluateurs.

Fiabilité intra observateur : C'est l'étendue avec laquelle un instrument produit des scores cohérents lorsqu'il est utilisé par le même évaluateur.

Fiabilité temporelle ou test-retest : C'est la stabilité d'une mesure au cours du temps. Le même évaluateur administre au sujet l'instrument de mesure à deux occasions avec un intervalle de temps durant lequel on espère qu'il n'y a pas de changement clinique (en général quelques jours à deux semaines).

Cohérence interne : Elle se réfère à l'homogénéité de la mesure en terme de groupements d'items du même type en unités.
Trois méthodes statistiques peuvent être utilisées pour analyser l'unidimensionnalité d'un instrument (58). Les deux premières sont classiques et la dernière l'est moins, elle est parfois utilisée dans des études :

- **L'analyse en composantes principales (ACP)** pour examiner la structure interne et l'unidimensionnalité de la mesure étudiée. L'ACP permet une réduction des données par construction de facteurs principaux représentant les dimensions de l'instrument. Elle étudie les distances entre les items et le positionnement de ces items par rapport à ces dimensions. Un instrument peut être considéré comme unidimensionnel si la variance expliquée par le premier facteur atteint un seuil entre 20 et 40 % de la variance totale et que le rapport de la première sur la deuxième valeur propre est supérieur à 2. Le seuil minimal adopté pour retenir une corrélation avec une composante principale est de 0,30.

- **Le coefficient alpha de Cronbach** mesure l'homogénéité de l'instrument et évalue la cohérence des items de l'instrument. Une fiabilité de 0,80 est nécessaire pour conclure à une bonne congruence des items.
- **La méthode basée sur le modèle de Rasch** permet de calculer l'échelle de difficulté des tests fonctionnels et vérifier l'unidimensionnalité de l'outil. Le postulat du modèle de Rash est l'existence d'une seule dimension, d'un seul concept sous-jacent à l'instrument. Le modèle de Rasch suppose, en effet, l'existence d'un « trait latent », c'est-à-dire d'une variable définie conjointement par les capacités des sujets et par les difficultés des items au regard du concept. Chaque item explore un certain niveau de difficulté du trait latent. Un individu d'une capacité donnée a d'autant plus de chance d'effectuer un test que sa capacité est supérieure à la difficulté de ce test, et d'autant moins de chance que sa capacité est inférieure à la difficulté du test considéré. Ce modèle repose sur des hypothèses jugées très fortes par de nombreux auteurs.

II.4.3 Validité.

C'est l'étendue avec laquelle un instrument mesure réellement ce qu'il est censé mesurer. Evaluer la validité implique de considérer le but de l'instrument, le type de validité et la qualité de la recherche.
Il existe quatre types de validité : de contenu, de construction, de critères et la sensibilité au changement. La validité de contenu est essentielle pour tous les types d'instruments. Les instruments évaluatifs et descriptifs requièrent une validité de construction. La validité de critères est essentielle pour les instruments prédictifs alors que la sensibilité au changement est importante pour les instruments évaluatifs.
Les résultats sont là aussi exprimés en coefficients de corrélation.

Validité de contenu : Elle est en rapport avec la compréhension de l'instrument. Un instrument qui a une bonne validité de contenu couvre le domaine pour lequel il a été construit. Cette validité est essentielle pour tous les types d'instruments mais il s'agit d'une validité basique qui doit être associée aux autres validités. Pour évaluer cette validité, on fait appel à des jugements d'experts ou bien à des méthodes statistiques.

Validité de construction : C'est l'étendue avec laquelle un instrument mesure ce qui était attendu théoriquement lors de la construction de l'instrument. En d'autres termes, on vérifie que les résultats obtenus sont bien en accord avec les hypothèses théoriques émises lors de la construction de l'instrument. Pour cela, des hypothèses variées doivent être examinées

Validité de critère : On compare la mesure à un « gold standard » c'est-à-dire à une mesure construite de la même façon et qui a démontrée une fiabilité et une validité. On peut aussi comparer à un critère objectif extérieur. Cette validité n'est pas essentielle pour un instrument prédictif.

Sensibilité au changement: C'est le degré avec lequel un instrument peut détecter un vrai changement. Elle peut être évaluée : soit en mesurant le changement au cours du temps de deux groupes, l'un sachant qu'il y a un changement et l'autre ne le sachant pas; soit en étudiant la capacité de l'instrument de mesurer les résultats d'un traitement connu pour être efficace.

II.4.4 Les normes :

Les normes sont essentielles pour les instruments prédictifs et descriptifs. Les normes d'un instrument sont significatives et faciles à utiliser sur le plan clinique, seulement si la fiabilité et la validité de l'instrument de mesure ont été établies.
Une description de la méthode et de l'échantillon utilisés pour établir les normes est utile pour pouvoir comparer avec d'autres populations et d'autres protocoles de tests.

II.5. Perspective du patient :

Est-ce que l'instrument inclus la perspective du patient ? Ceci comporte de nombreux avantages :
1) les besoins, les désirs et l'impact de l'environnement du patient peuvent être étudiés.
2) on peut évaluer la différence entre ce que le patient peut faire et ce qu'il fait vraiment.
3) cela doit faciliter une collaboration patient/thérapeute pour la réadaptation.
Si un instrument n'a pas ce critère il ne doit pas pour autant être écarté et on peut toujours utiliser en association un autre instrument qui prend en compte ce critère.

Conclusion :

Un tableau résume les différentes qualités d'un instrument d'évaluation en général et de la main en particulier. Ce tableau est tiré de l'article de Rudman(77).

Catégorie 1 : Utilité clinique
Applicabilité clinique
Spécificité
Disponibilité
Temps demandé
Acceptabilité par le patient

Catégorie 2 : Standardisation
Instructions
Equipement nécessaire

Catégorie 3 : Buts /intentions du test
Importance ++++

Descriptif	**Prédictif**	**Evaluatif**

Catégorie 4 : propriétés psychométriques

A) Construction des items

Descriptifs	Prédictifs	Evaluatifs

B) Fiabilité

Inter observateurs	Inter observateurs	Inter observateurs
Intra observateur	Intra observateur	Intra observateur
Test-retest	Test-retest	Test-retest
Cohérence interne	Cohérence interne	Cohérence interne

C) Validité

Contenu	Contenu	Contenu
Construction	Construction	Construction
Critères	Critères	Critères
Sensibilité chgt	Sensibilité chgt	Sensibilité chgt

D) Normes

Disponibilité	Disponibilité	Disponibilité
Qualité	Qualité	Qualité

Catégorie 5: Perspectives du patient

Tableau n°1: Etapes utiles pour sélectionner un instrument d'évaluation de la main. En rouge les qualités indispensables requises pour un instrument dans chaque catégorie (évaluatif, descriptif et prédictif). D'après Rudman (77).

III. TESTS REPETITIFS DE DEXTERITE.

L'évaluation de la dextérité est une partie importante de l'évaluation fonctionnelle de la main, mais elle ne saurait la résumer.
Elle fait appel aux différentes préhensions possibles avec la main, mais elle dépend aussi de la motricité, de la sensibilité et des performances corticales. Dans une première partie je reviendrai sur différentes définitions concernant la dextérité et les préhensions. Dans une deuxième partie je ferais un rapide bilan des tests classiques de dextérité utilisés de manière régulière par les thérapeutes de la main. Une troisième partie sera consacrée à deux tests récents développés pour l'étude de la dextérité. Enfin je conclurai sur l'intérêt de ces tests en pratique courante.

III.1 DEFINITIONS DE LA DEXTERITE ET DES PREHENSIONS.

LA DEXTERITE : d'après le petit Robert, vient du latin dexteritas, de dexter (dextre) qui est la main droite. Elle est définie comme l'adresse des mains, la délicatesse, l'aisance dans l'exécution de quelque chose.
Bear (9), la définit comme l'habileté et la facilité à utiliser les mains.
Pour Xenard (96), la dextérité réfère globalement à l'habileté à utiliser ses mains ou la capacité à utiliser volontairement des mouvements fins afin de manipuler adéquatement de petits objets dans une tâche spécifique. La précision et la vitesse sont les paramètres de mesure de la dextérité (1).
La dextérité est directement dépendante de plusieurs facteurs (23,96) :

- Des préhensions (cf. infra)
- De la coordination motrice qui est la capacité du système nerveux d'organiser des activités musculaires en vue de la réalisation correcte du mouvement.
- Des habiletés acquises par l'entraînement consécutif à la manipulation d'objets.
- De l'âge, du sexe.
- De la sensibilité.
- Des variations anthropométriques.

LA PREHENSION : d'après le Petit Robert, du latin prehendere (saisir), c'est la faculté de saisir avec un organe approprié.
Delprat (25), la définit comme l'action de prendre, d'agir sur ou avec un objet, mais aussi de l'explorer. Ceci se fait chez l'homme grâce à l'existence d'une opposition du pouce presque complète.
La préhension est le résultat de processus successifs d'intention, de représentation, de planification, aboutissant à l'exécution (25).
La première phase est la préparation dépendante de la vue, des expériences antérieures mémorisées pour l'identification des paramètres de l'objet. La deuxième phase est l'approche de l'objet qui nécessite un ajustement postural. La troisième phase est la saisie proprement dite avec un temps de solidarisation, un temps d'exploration pendant lequel se réalise le projet et un temps de dégagement. Ces stratégies d'approche et de préhension varient selon l'usage qui doit être fait de l'objet et selon les individus. En pathologie, il existe de grandes possibilités d'adaptation, de suppléance et de compensation.

Différents types de préhensions sont décrits avec des exemples précis. Il faut noter qu'il existe plusieurs terminologies différentes qui se recoupent facilement (23,25).

- La prise palmaire sphérique ou prise pollicidigitale en abduction ou multi pulpaire : prendre une balle de tennis.
- La prise palmaire cylindrique : prendre une bouteille.
- La prise en levier ou digitothénarienne , qui est une prise de force : utiliser un marteau.
- La prise en crochet : porter un sac à main.
- La prise de soutien : porter un plateau.
- La pince tri digitale : pouce-index-majeur : utiliser un crayon.
- La pince en opposition pulpo-latérale ou sub termino latérale avec opposition incomplète du pouce : tourner une clé.
- La pince interdigitale qui nécessite peu de force et est une prise d'attente : fumer une cigarette.
- La pince unguéale ou pince termino terminale ou pince ronde avec opposition complète du pouce : ramasser une aiguille.
- La pince pulpaire ou subtermino-subterminale ou pince longue : prendre une feuille de papier.

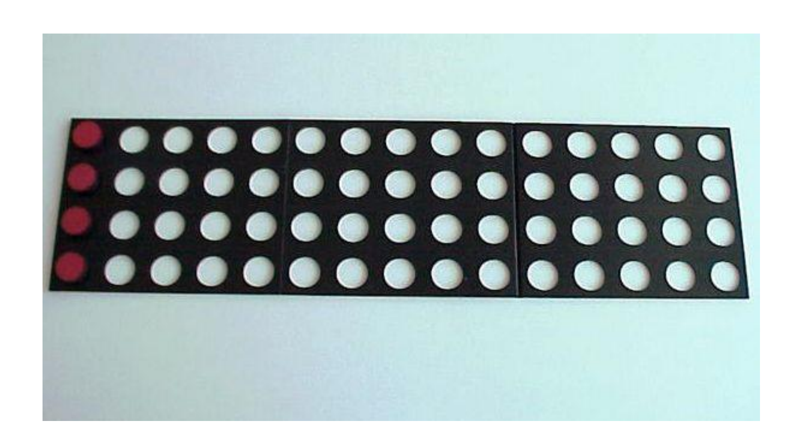
Parfois, il existe des associations de deux types de préhension sur une même main. De même, trois types de préhensions bi manuelles existent, la prise à deux mains (addition des deux mains) pour un usage de force (utilisation d'une pioche), les préhensions de l'une et l'autre main avec une action différente pour chaque main sur l'objet saisi (ouvrir une bouteille) et enfin les prises avec les deux mains pour saisir des objets volumineux (porter une caisse).
Lors de l'étude des préhensions, on étudie bien sur la façon dont la préhension se fait mais aussi s'il existe des compensations, des gestes douloureux. La cotation se fait le plus souvent en normal, perturbé et pas réalisable (23).
Nous verrons que les tests d'évaluation fonctionnelle de la main, en particulier les tests de la vie quotidienne, reposent sur l'exploration de ces différents types de préhensions.

III.2. LES TESTS « CLASSIQUES » REPETITIFS DE DEXTERITE.

La plupart de ces tests sont anciens et ont été mis au point pour la sélection des ouvriers à l'embauche dans des emplois répétitifs en usine. Certains de ces tests sont utilisés en pratique courante par les ergothérapeutes et sont de bons tests pour se faire une première idée de la dextérité. Voici les principaux tests que l'on peut retenir (9, 23,25).

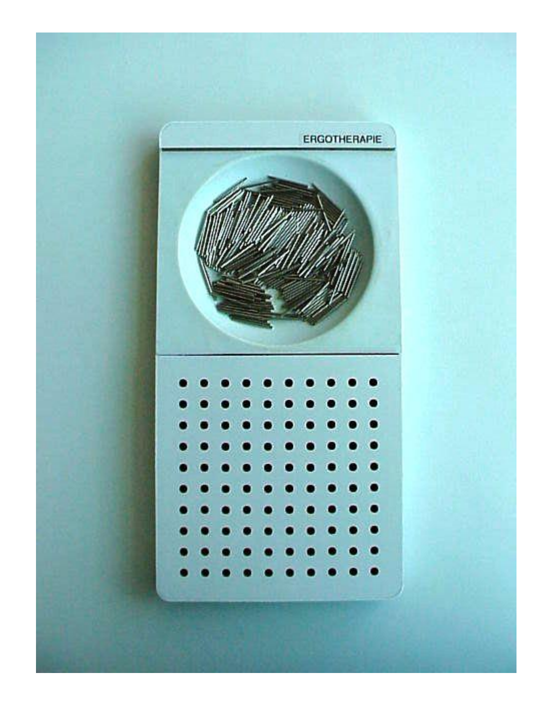


III.2.1. Test de dextérité manuelle de Minnesota.



Il date des années 1940 et il teste la coordination grossière et la dextérité. Le protocole est bien codifié et standardisé. L'examinateur doit expliquer et démontrer au patient la façon de procéder. Le sujet est en position debout, et doit placer le plus rapidement possible des pions d'un diamètre de 37 mm dans des trous situés sur une planche posée sur une table devant la personne (placing test). Cette première partie teste la main dominante. Dans une deuxième partie, la personne doit retourner ces pions à l'aide des deux mains (turning test). Le temps en secondes que met la personne pour effectuer ces tâches est mesuré et comparé à une échelle en percentile. Deux, trois ou quatre essais doivent être exécutés.
Ce test ne montre pas une vue convenable de la fonction de la main mais donne des informations suffisantes pour prédire une performance au travail. Il exige une bonne organisation motrice, une grande dextérité, la coordination œil main, l'endurance du geste et une bonne mémoire (24).L'avantage est qu'il est simple, reproductible et standardisé.
La fiabilité test retest a été évaluée sur deux essais entre 0.87 à 0.95 et pour quatre essais entre 0.93 et 0.98. La validité du test a été étudiée chez des ouvriers de manufacture de papier par Jurgensen en 1943. Trois évaluateurs indépendants ont notés chaque travailleur sur une échelle de 5. Les résultats ont été convertis en T score et la somme des trois T scores a été utilisée comme le critère de succès. La fiabilité de la mesure était de 0.75. Quatre tests (placing, turning, une main turning et placing, deux mains turning et placing) ont été administrés. La corrélation entre les quatre tests et le critère de succès était de 0.32 à 0.57.
Le Minnesota a été comparé à un autre test le Pensylvania Bi-manual work sample, chez 477 patients. Les corrélations étaient de 0.46 pour le placing test et 0.40 pour le turning test.

III.2.2. Test de dextérité digitale selon O'Connor.



Il date de la fin des années 1920. Il teste la dextérité fine digitale de la main dominante. La personne testée est en position assise et doit placer des petites tiges métalliques de 25 mm de longueur dans cent trous répartis sur dix rangées, sur une planche placée devant elle. La personne doit prendre chaque fois trois tiges en même temps aussi vite que possible. Le temps que met la personne est mesuré en secondes et comparé à une échelle percentile. Comme le précédent, il est simple, mais il ne teste que la main dominante. Je n'ai pas trouvé de données sur la standardisation, la fiabilité, la reproductibilité.



III.2.3. Test de dextérité fine et de rapidité de Purdue ou Purdue pegboard test.



C’est le plus connu de ces différents tests et il explore la dextérité fine. Il a été mis au point en 1948. Il existe des normes pour les étudiants hommes et femmes, les ouvriers de l’industrie et les vétérans masculins de l’armée.
Il se pratique en position assise. Le patient doit placer le maximum de tiges métalliques de 25 mm de long dans une planche placée sur une table devant lui dans un temps donné. L’évaluation est faite avec la main dominante, puis la main non dominante puis avec les deux mains ensemble. Il requiert une bonne coordination œil main. Ensuite la personne doit constituer des assemblages en utilisant simultanément les deux membres supérieurs pour placer dans un ordre donné le maximum de tiges métalliques, de rondelles et de cylindres. Les deux mains travaillent tout le temps ensemble.
La fiabilité test-retest est bonne de 0.60 à 0.75 voire à 0.91 (moyenne de trois études). Elle a été étudiée dans plusieurs groupes de professions.
Pour la validité, de nombreuses études ont été faites étudiant diverses professions. Pour ce faire, on rapporte le score du test aux véritables performances d’un travail spécifique pour lequel le test a été utilisé comme outil de sélection.
Chez des stagiaires en école de commerce la corrélation était de 0.26 ($p<0.01$) entre la somme des trois tests (main dominante, non dominante et les deux ensemble) et le critère retenu (évaluation du temps), mais il n’existait pas de corrélation si le critère était l’évaluation de la qualité du travail. Chez des opérateurs de machine, le critère choisi étant la production, il existait une corrélation de 0.61 ($p< 0.01$) avec les tests d’assemblage. Chez des emballeurs, le critère choisi était l’évaluation par le superviseur et la corrélation était de 0.47 ($p<0.01$) avec les tests d’assemblage.

III.2.4. Test d'aptitude à la manipulation selon Roeder.



Le patient est assis et doit visser le maximum de petites tiges métalliques de 20 mm de long dans une planche placée sur une table devant lui et visser un écrou borgne sur chaque tige métallique dans un temps donné avec la main dominante. Dans un deuxième temps, le patient doit glisser autant de rondelles et de gros écrous que possible, en les alternant sur une barre en T avec la main dominante, la main non dominante et les deux mains ensemble.
Le résultat est exprimé en nombre de tiges, écrous et rondelles comparé à un percentile.
La fiabilité test retest a été évaluée à 0.92.
La validité a été étudiée dans l'industrie électronique dans deux études (corrélation à 0.49 et à 0.48) en utilisant comme critère l'évaluation du superviseur.

Les intérêts principaux de ces quatre tests sont :

- La facilité et la rapidité d'utilisation.
- La bonne reproductibilité, la standardisation sauf pour le O'connor ou les données sont inconnues.

Les désavantages sont :

- L'absence de proximité de ces tests avec les activités de la vie quotidienne.
- Le fait que ces tests aient été développés pour sélectionner des travailleurs à l'embauche pour des activités répétitives et non pour tester des malades dans une situation donnée.
- Le temps comme principale mesure.
- L'absence de validation de tous les critères d'un bon instrument descriptif et évaluatif (cf tableau infra).

PROPRIETES	MINNESOTA	O'CONNOR	PURDUE	ROEDER
Utilité clinique	+	+	+	+
Standardisation	+	+	+	+
Buts du test	**Descriptif Evaluatif**	**Descriptif Evaluatif**	**Descriptif Evaluatif**	**Descriptif Evaluatif**
Fiabilité				
Inter-obs	?	?	?	?
Intra-obs	?	?	?	?
Test retest	0.87 à 0.98	?	0.60 à 0.91	0.92
Cohérence int	?	?	?	?
Validité				
Contenu	A priori oui	A priori oui	A priori oui	A priori oui
Construction	0.32 à 0.57	?	0.26 à 0.61	0.49
Critères	Pas utile ici	Pas utile ici	Pas utile ici	Pas utile ici
Sensibilité chgt	?	?	?	?
Normes	+	+	+	+
Perspectives du patient	Non	Non	Non	Non

Tableau n°2 : Propriétés des quatre tests de dextérité « classique ».

Dans un instrument évaluatif (cf supra), les critères psychométriques importants sont la fiabilité inter, intra observateur et test retest, la validité de contenu et de construction et la sensibilité au changement. Pour trois d'entre eux seules les données de la fiabilité test retest et de la validité de construction sont disponibles. La validité de contenu est a priori bonne. Par contre la sensibilité au changement n'est jamais disponible ainsi que la cohérence interne qui est importante dans un outil descriptif. Il est vrai qu'il s'agit de tests très anciens dont les critères de construction à l'époque n'étaient pas aussi stricts qu'à l'heure actuelle. Néanmoins, ces tests sont utilisés dans notre expérience pour un premier bilan de la dextérité et permettent de mettre en évidence des incapacités qui vont nécessiter une prise en charge thérapeutique ciblée. En pratique le Purdue et le Minnesota sont surtout utilisés. Ces tests doivent être complétés par d'autres tests que nous reverrons plus loin. En l'absence de données plus précises le test de O'Connor devrait être abandonné.
Le tableau ci-dessous résume ces quatre tests.

MINNESOTA	O'CONNOR	PURDUE	ROEDER
Main dominante Puis bi manuel	Main dominante	Main dominante Puis non dominante Puis bi manuel	Main dominante Puis non dominante Puis bi manuel
Temps en secondes Comparaison à percentile	Temps en secondes Comparaison à percentile	Temps en secondes Comparaison à percentile	Nombre écrous, vis, rondelles Comparaison à percentile
Coordination et dextérité grossières.	Dextérité digitale	Dextérité fine et rapidité	Manipulation fine

Tableau n°3 : Caractéristiques des quatre tests de dextérité « classique ».

III.2.5. Autres tests.

De nombreux autres tests existent et ne seront pas développés de manière extensive (9,25).

Le Crawford small parts dexterity test: étudie la coordination motrice fine en utilisant une petite pince de brucelles (pince fine à ressort) et un petit tournevis pour manipuler des épingles, des colliers et des petites vis. La fiabilité test retest et la validité de contenu sont bonnes et la validité prédictive non étudiée.

Le Bennett hand-tool dexterity : utilisé pour mesurer la capacité à utiliser des outils mécaniques ordinaires. Les outils sont utilisés pour attacher et défaire des écrous et des boulons. Le test est bi manuel. Ce test étudie surtout la partie physique des activités. Il est non standardisé mais la fiabilité test retest est bonne (0.81 à 0.88).

Le box and block test of manual dexterity : évalue des déficits sévères de la dextérité. Le patient transfère le plus possible de cubes de 2,5 cm de côté d'une boîte à une autre en une minute. Le nombre de cubes transportés donne le score rapporté sur 10. Il est utilisé surtout en neurologie. La fiabilité test retest est de 0.937 et 0.976, la fiabilité inter observateurs est de 0.999 pour la main gauche et de 1.0 pour la main droite (68).

Le nine hole peg test of finger dexterity: est un test qu'il ne faut pas utiliser seul mais en association avec d'autres tests comme le Purdue. La fiabilité inter observateurs est élevée 0.97 pour la main droite et 0.99 pour la main gauche, par contre la fiabilité test retest est mauvaise 0.69 pour la main droite et 0.43 pour la main gauche (68). Il existe des normes par sexe et âge. Il n'a que peu d'intérêt par rapport au Purdue.

On peut encore citer le Rosenbuch test (43) qui explore la dextérité fine manuelle unilatérale et interdigitale. Sa cotation dépend de la rapidité et de la qualité d'exécution du geste. Il semble être valide et fiable.

III.3. LES TESTS RECENTS DE DEXTERITE :

De façon récente c'est-à-dire depuis les années 1990 de nouveaux tests de dextérité ont été développés. Ils répondent à des critères plus stricts de développement mais restent encore à l'heure actuelle difficiles à utiliser en pratique courante.Il s'agit du NK hand dexterity test (NKHDT) et du functional dexterity test (FDT):

III.3.1. Le NK hand dexterity test: NKHDT (6, 61, 87).

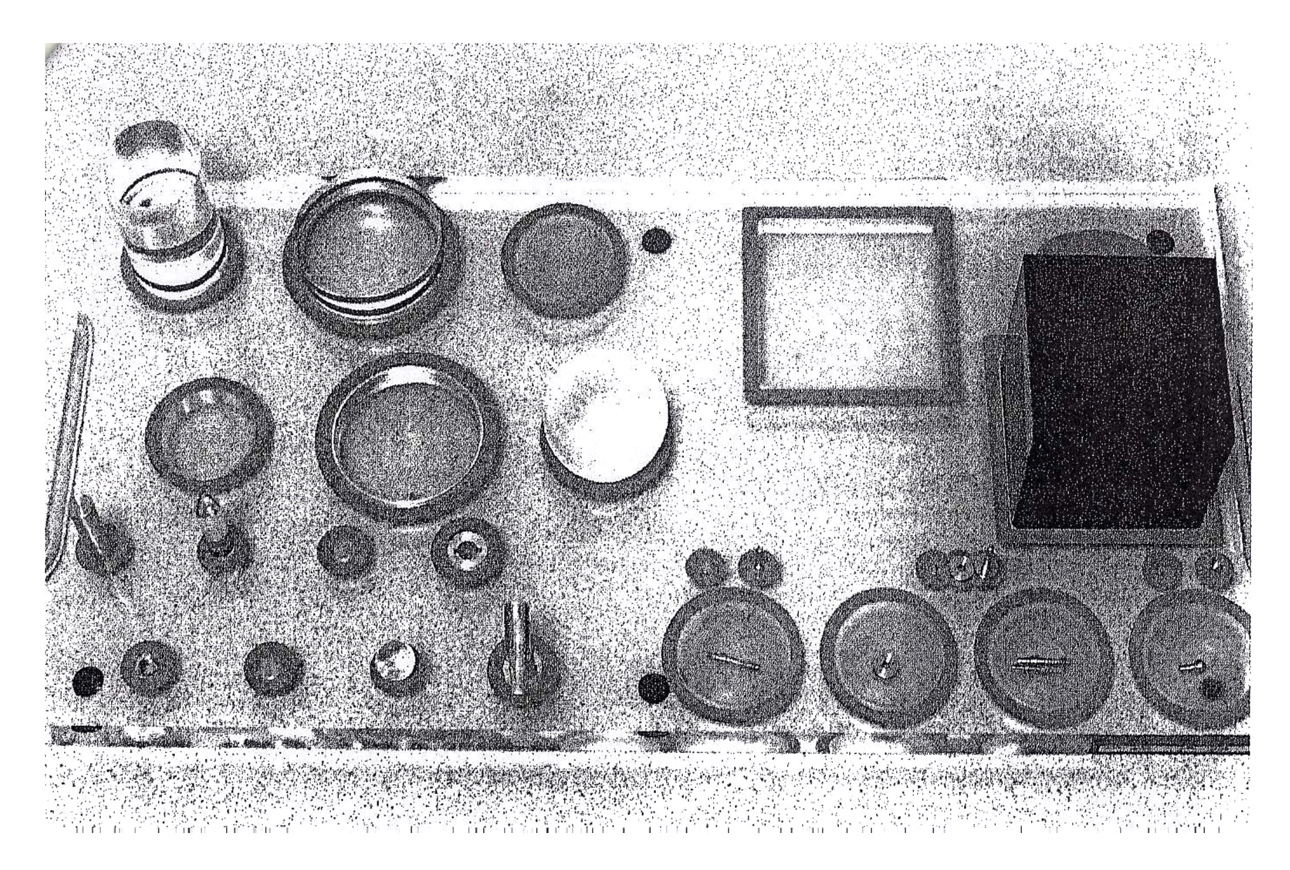
En 1994, la firme NK a mis au point un test de dextérité de la main ayant la particularité d'être couplé à un ordinateur et un logiciel.

Amadio en 1996 (6) avait montré que chez 21 patients ayant une fracture de Pouteau-Colles, le NKHDT avait une SRM (réponse standardisée moyenne cf p75) de 0. 83 et était donc à même de détecter un changement clinique. Dans cette étude, le NKHDT était plus sensible que la force au Jamar, au pinch et que la sensibilité. Il restait à déterminer la fiabilité du test et son potentiel d'utilisation. Il faut rappeler ici que cet appareil est très coûteux (1700 dollars pour le clavier seul sans compter le logiciel et le PC).

Turgeon et collaborateurs en 1999 ont mis sur pied une étude dont le but était d'étudier la cohérence des résultats dans deux situations d'utilisation du NKHDT (87).

- *Matériel et méthodes* :

  Le clavier de dextérité est composé de petits, moyens et gros objets à manipuler. Ce clavier est relié à un ordinateur et un logiciel calcule automatiquement le temps nécessaire à la manipulation des objets. Le logiciel contient des normes pré-établies des différents tests.



Les sujets étudiés étaient 37 sujets sains, 24 femmes et 13 hommes, d'age moyen 29 ans, dont 34 sur 37 étaient droitiers.
La procédure du test n'est pas standardisée par le fabricant. Dans cette étude chaque groupe d'objets a été testé séparément (petits, moyens et gros). Des instructions précises sont données sur la façon de bouger les objets en utilisant une seule main par essai. Une fois que les objets sont tous déplacés il faut les remettre à leurs places initiales. La procédure est répétée deux fois de chaque côté. Le temps pour chaque sous-test est calculé par l'ordinateur.
20 sujets ont été testé à nouveau une à cinq heures après et les 17 autres 3 à 14 jours après.
L'analyse statistique de la fiabilité a reposé sur le calcul du coefficient de corrélation intra classe (ICC) entre les essais pour chacun des deux tests pour chaque sujet, et lors des retests entre les scores des premiers, deuxième et la moyenne des deux. Une excellente corrélation était associée à un ICC > 0.75, une acceptable à bonne si ICC entre 0.40 et 0.75 et une mauvaise si ICC< 0.40.

- *Résultats :*

La majorité des sujets se plaignent de fatigue du bras ou de l'avant-bras en manipulant les gros objets. Les objets sont considérés comme trop durs et le test trop long. Le temps nécessaire pour faire le test a été de 20 à 25 minutes la première fois et de 15 à 20 minutes la deuxième.
La main dominante effectue le test plus rapidement que la non dominante.
L'ICC de la fiabilité entre les essais au cours du même test : sur les 12 valeurs 6 sont entre 0.40 et 0.75 et 6 sont supérieures à 0.75. La fiabilité est excellente pour la main dominante et acceptable pour la main non dominante.
ICC de la fiabilité test-retest : tous les ICC pour les objets petits et moyens étaient acceptables entre 0.40 et 0. 75. Pour les gros objets la majorité des ICC étaient supérieurs à 0.75. La main dominante était là aussi plus performante.

- *Discussion :*

Dans l'évaluation clinique, les thérapeutes savent que les sujets peuvent faire lors des tests des variations dans le résultat même s'il n'y a pas de changement clinique. Des changements discrets de la dextérité ne peuvent être attribués aux effets du traitement. Pour minimiser cet effet, les cliniciens ont besoin de faire la moyenne de ces tests pour établir une ligne de base stable. Dans cette étude, les scores de la main dominante sont plus fiables ce qui reflète le fait qu'elle a une plus grande dextérité et que des actes de maladresse sont moins susceptibles d'affecter les scores de la main dominante. L'on voit aussi que l'effet d'apprentissage améliore les résultats et contribue aux variations entre les scores. D'autres études seraient utiles pour déterminer le nombre optimal de répétitions et établir un protocole pratique qui contrebalance les effets de l'apprentissage et de la fatigue.
Avec les autres tests de dextérité, les ICC sont similaires, mais en raison des différences de méthodes pour reporter la fiabilité, la comparaison directe avec ces études pose problème.
Ce test étant récent il reste à préciser le protocole exact du test et à le rendre encore plus fiable. Des modifications doivent aussi être apportées au tableau de dextérité en ce qui concerne les objets et le matériau utilisé.

- *Avantages du test :*

Teste une large variété de mouvements fins et grossiers.
L'enregistrement des données est faite par l'ordinateur.
Il existe des normes fournies avec le logiciel.

Le NHKDT a montré dans une étude qu'il était sensible (6).

- *Désavantages du test :*

Le tableau est volumineux.
Le coût.
Une fiabilité uniforme indiscutable n'est pas encore prouvée.
Il est prématuré de faire confiance aux valeurs normales fournies par le logiciel tant qu'un protocole standardisé et très fiable n'aura pas été publié.

McDermid en 2001 a publié la deuxième étude disponible pour le NKHDT (61). Cette étude s'intéresse à la validité du NKHDT. A la différence de l'étude précédente, McDermid accepte d'emblée les valeurs normales fournies par le logiciel. Pour la validité de critères, il utilise comme comparatif le Jebsen hand function test (JHFT), qui est un test utilisé ici comme le « gold standard ». Nous reverrons en détail ce test au chapitre suivant. La littérature suggère que ce test est valide comme indicateur de la dextérité de la main
Les trois buts de l'étude sur la validité du NKHDT étaient : 1) comparer les résultats de ce test avec ceux du JHFT ; 2) déterminer l'importance que chaque test de dextérité a dans la capacité fonctionnelle du patient ; 3) déterminer la capacité de chaque test de discriminer les patients ayant ou non une pathologie du membre supérieur.

- *Matériel et méthodes :*

Il s'agit d'une étude de validité croisée. Les scores du NKHDT ont été corrélés aux scores du JHFT et à l'incapacité rapportée par les patients, cela pour établir la validité de critères.
La validité de construction a été testée en comparant les scores de sujets sains et de sujets avec pathologie du membre supérieur.
Cinquante sujets, vingt-cinq hommes et vingt cinq femmes d'âge moyen 40 ans, droitiers dans quarante-quatre cas ont été recrutés. Seuls dix sujets étaient sains, les quarante autres avaient des pathologies variées de la main. L'ordre des tests était randomisé et tous les tests ont été pratiqués par le même kinésithérapeute.
Pour le JHFT, le protocole suivi était celui de l'auteur (cf chapitre suivant), chronométré en secondes.
Pour le NKHDT, un test d'essai a eu lieu suivi de deux tests de mesure. La moyenne des deux tests a été utilisée ensuite. Le protocole a été le même que dans l'étude de Turgeon. Pour les sujets sains les tests sont faits avec la main dominante et pour les sujets malades avec la main atteinte.
Pour l'autoévaluation par le patient, la sous échelle activités usuelles du questionnaire PRWE (patient rating wrist /hand evaluation) (60) a été utilisée. Le sujet cote son niveau de difficulté sur une échelle de 1 à 10 pour quatre activités : soins personnels, activités ménagères, travail et activités de loisir. Cette sous échelle a montré une fiabilité test-retest de 0.88 (66).
L'analyse statistique a comparé chaque score obtenu aux tests aux scores normaux du JHFT et du NKHDT en corrigeant en fonction de l'âge, du sexe et du côté dominant. Le coefficient de corrélation de Pearson r entre les tests et l'autoévaluation du patient a été calculé. La corrélation était pauvre si $r < 0.40$, modérée si r entre 0.40 et 0.75 et haute si $r > 0.75$. Une analyse uni variée de la variance (ANOVAs) a été conduite pour détecter une différence significative entre les sujets sains et malades.

- *Résultats :*

Les coefficients de corrélation entre JHFT et NKHDT sont entre 0.34 et 0.87. Les sous tests avec des objets de même taille sont mieux corrélés que pour des objets de tailles différentes. Les deux tests sont modérément corrélés avec l'autoévaluation du patient. La force de ces

corrélations était similaire pour les deux tests. Les sous tests du NKHDT et du JHFT étaient légèrement plus fortement corrélés entre eux qu'avec le questionnaire.
Comme attendu, les personnes malades ont eu plus de difficultés à faire les tests que les personnes saines. Il y avait une différence statistiquement significative ($p < 0.01$) entre les sujets normaux et les malades et ceci pour les deux tests et le questionnaire.

- *Discussion :*

Cette étude confirme la validité du NKHDT. La validité concurrente avec le JHFT, basée sur l'hypothèse que ces deux tests sont valides, chronométrés et faits d'items qui étudient la dextérité manuelle est confirmée. Comme les deux tests sont composés de différentes tâches, les corrélations sont modérées, mais les sous tests qui se ressemblent ont une corrélation plus forte.
La dextérité n'est qu'une partie de la fonction et par conséquent les corrélations modérées entre la dextérité et l'autoévaluation du patient sont considérées comme reflétant une importante relation entre ces deux variables. Les deux tests sont corrélés de façon similaire à l'évaluation du patient ce qui confirme aussi la validité du NKHDT. Les deux tests arrivent à discriminer les patients sains et malades ce qui est aussi un argument pour la validité.
La principale limitation de cette étude vient de l'utilisation du JHFT comme comparatif. Ce test est considéré comme valide dans son contenu, mais il manque de standardisation et d'explications de détails pour une précision optimale et ses propriétés psychométriques sont douteuses (cf. chapitre suivant).

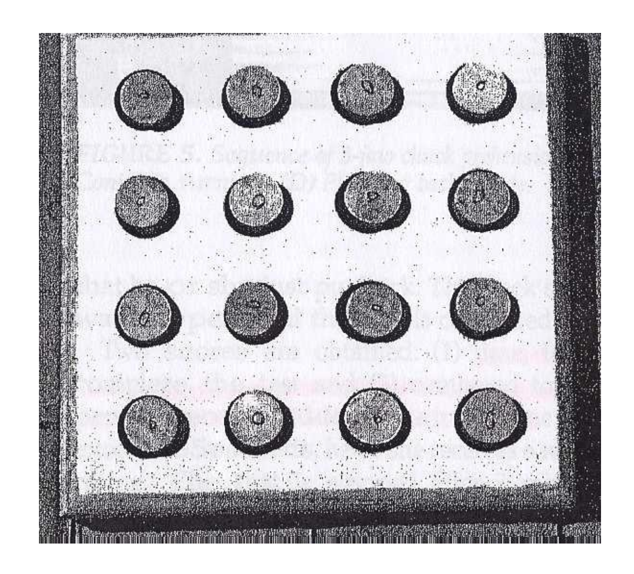
En conclusion de ces deux études, on peut dire que le NKHDT peut être utilisé pour évaluer la dextérité de la main car il est fiable, sensible et valide. Le tableau n°4 résume les propriétés de ce test. Néanmoins, il n'est pas facile d'utilisation (encombrement), il est coûteux et n'apporte à mon avis pas grand-chose de plus que les tests plus simples dans une évaluation quotidienne de la dextérité. Son principal intérêt réside dans le fait qu'il a été bien étudié sur le plan statistique et qu'il est automatisé.

III.3.2. Le Functional Dexterity test (FDT):

Ce test a été développé sur vingt ans de 1982 jusqu'à sa publication in extenso en 2003 (1). Il a été élaboré pour combiner trois composants de la dextérité :
1) la manipulation (dynamique, de précision).
2) le temps et la précision (vitesse).
3) la préhension tri digitale (pouce-index-médius).
Le but du FDT est de mesurer la capacité du sujet à faire une prise tri digitale dynamique chronométrée en manipulant des chevilles sur un tableau. La prise tri digitale est appelée main primaire et différencie l'homme du singe. Elle est très importante dans de nombreuses activités de précision (écrire, boutonner, attacher).

- Développement de l'instrument :

-Buts : aide rapide pour le clinicien pour évaluer la préhension tri digitale.
-Equipement : tableau en bois carré avec seize grosses chevilles.
-Instructions pour administrer le test : le patient est assis. L'examinateur demande au sujet de tourner toutes les chevilles avec la main saine en partant du haut à la partie opposée du tableau et en avançant en zigzag. Si une cheville tombe le temps est stoppé et une pénalité de dix secondes est donnée. Le patient a droit à un essai libre. Le test est réalisé deux fois avec la main saine puis la main lésée.



- Calcul du score :

Le temps est calculé en secondes. Deux scores sont obtenus, le temps pour finir le test et le temps combiné avec les pénalités éventuelles. Si le temps excède 55 secondes le score est dit non fonctionnel. Le score final est utilisé pour déterminer le niveau fonctionnel en fonction du temps : main fonctionnelle, modérément fonctionnelle, fonctionnelle un minimum ou non fonctionnelle. Un tableau est disponible pour le calcul.

- Estimation de la fiabilité :

La fiabilité inter observateurs a été étudiée par le coefficient de corrélation intra classe (ICC). Ce coefficient était excellent toujours supérieur à 0.99 pour les scores chronométrés sans pénalités, que ce soit pour la main lésée ou la main saine. Pour les scores avec pénalités, pour les mains lésées les ICC sont supérieurs à 0.82 et pour les mains saines à 0.67.
La fiabilité intra observateur étudiée aussi par l'ICC montre pour la main lésée, avec ou sans pénalités un ICC supérieur à 0.9, et pour la main saine avec pénalités un ICC supérieur à 0.72. La fiabilité est donc excellente pour les scores bruts, et bonne pour les scores avec pénalités.
Il existe un effet d'apprentissage significatif qui a été réduit en demandant aux patients de faire un essai d'apprentissage complet avec la main saine. Ceci a permis d'améliorer les ICC des scores sans et avec pénalités jusqu'à 0.90.

- Validité du test :

Pour la mettre en évidence, il a été recherché une relation entre les scores au FDT et des activités de la vie quotidienne qui nécessitent l'utilisation de la pince tri digitale. 46 patients ont été testés, 35 hommes et 11 femmes. Les scores au FDT ont été comparés avec la capacité des patients de faire quatre activités : boutonner, faire un lacet, visser un écrou et un boulon, nouer un fil. Ces activités ont été classées en quatre classes comme déjà indiqué ci-dessus. La corrélation entre les 4 tâches et le FDT était de -0.64 ($p < 0.001$) pour la main lésée non dominante et de -0.63 ($p < 0.02$) pour la main lésée dominante. Ceci indique que plus le patient est capable de faire les tâches fonctionnelles moins il met du temps pour faire le FDT.
Ceci prouve que les scores de la main lésée au FDT sont associés à la performance pour réaliser les tests de la vie quotidienne.

Une autre étude a comparé le FDT au JHFT chez 21 patients. Le coefficient de corrélation de Pearson entre JHFT et FDT était de 0.52 ($p<0.0015$) pour la main dominante (lésée ou non) mais non significatif pour la main non dominante. Quand on prend les sous tests du JHFT qui utilise la pince tri digitale le coefficient de corrélation est supérieur à 0.47 pour la main dominante (lésée ou non) et supérieur à 0.60 si l'on prend juste la main lésée.
Il existe donc une corrélation positive entre les scores au FDT et la capacité à faire des tâches qui nécessitent la pince tri digitale, en particulier pour la main dominante.

- Résultats des normes :

Des normes ont été établies pour différents groupes par âge, dominance et profession. Le temps pour réaliser le FDT augmente avec l'âge pour les deux mains.

En conclusion :
Le FDT a été bien étudié et le tableau ci-dessous montre ses principales caractéristiques. Il s'agit d'un outil fiable, valide, rapide à faire (15 secondes à 2 minutes) et qui étudie bien ce pour quoi il a été conçu. La sensibilité au changement n'a pas été étudiée. Une étude prospective à grande échelle est nécessaire pour confirmer les valeurs des normes avec des valeurs moyennes et des intervalles de confiance.
Ce test devrait encore faire l'objet d'une étude sur la validité de construction pour la confirmer. Ce test n'a pas été développé pour les patients avec lésions cérébrales qui pourraient entraîner des difficultés de compréhension du test. Il serait intéressant de le comparer avec le Minnesota dont le turning test semble assez proche du FDT.

PROPRIETES	**NKHDT**	**FDT**
Utilité clinique		
Applicabilité clinique	+	+
Spécificité	+/-	+/-
Disponibilité	+/- (coût)	+
Temps demandé	+	+
Acceptabilité	?	+
Standardisation	- par le fabricant + dans les deux études	+
Buts du test	Evaluatif descriptif	Evaluatif descriptif
Fiabilité		
Inter-observateurs	Test automatisé	>0.67 >0.82 > 0.99
Intra-observateur	Test automatisé	>0.72 >0. 9
Test retest	ICC entre 0.40 et 0.81	0.92
Cohérence interne	?	?
Validité		
Contenu	+	+
Construction	+	+pearson>0.6 avec JHFT
Critères	pearson entre 0.47 et 0.87	?
Sensibilité au changement	Oui SRM de 0.83	?
Normes	+ d'après le constructeur à valider d'après Turgeon et coll.	+/-
Perspectives du patient	non	non

Tableau n°4 : caractéristiques des tests NKHDT et FDT

III.4. CONCLUSIONS.

La dextérité n'est pas le seul déterminant de la fonction de la main. La douleur, la mobilité, la force contribuent aussi à la capacité fonctionnelle. Nous avons vu dans ce chapitre que les différents tests classiques de dextérité ont le temps comme principal indicateur de capacité. McDermid (61) avance que le temps mis pour réaliser des tâches ne serait pas un indicateur valide de l'incapacité fonctionnelle ou du handicap. En effet, les sujets évaluent leur capacité à faire des activités variées de la vie quotidienne, non pas sur le temps mais sur leur possibilité à faire ces tâches avec succès. Le temps mis pour faire certaines tâches évalue sûrement une capacité physique, mais ne tient pas compte de la qualité avec laquelle est faite cette tâche. Or, il paraît très important dans l'évaluation fonctionnelle de la main de s'attacher à la qualité des tâches accomplies et aux éventuelles stratégies et compensations mises en œuvre par les patients pour palier une incapacité. Ces tests de dextérité ont été développés essentiellement dans un but de sélection à l'embauche de sujets sains et non pas d'évaluation fonctionnelle de patients porteurs de pathologies manuelles. Même s'ils apportent des informations pertinentes sur l'efficacité du patient à manipuler des objets, ils ne peuvent résumer une bonne évaluation fonctionnelle de la main. D'autres tests plus proches de la réalité et développés pour l'évaluation de patients seraient utiles. Ceci fait l'objet du prochain chapitre où nous allons voir qu'il existe de nombreux tests dit de vie quotidienne dont certains peuvent compléter utilement les tests classiques.

IV. LES TESTS DE DEXTERITE DE LA VIE COURANTE.

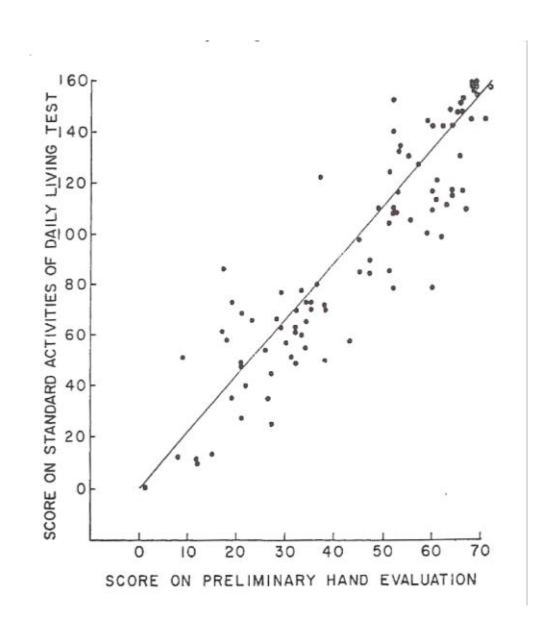
IV.1 Définitions et qualités requises.

C'est en 1965 que le premier de ces tests a été publié par Caroll (19). L'idée était de faire un test simple à utiliser, rapide et surtout dont les items auraient une relation directe avec la vie quotidienne des patients. Dans la plupart de ces tests le temps seul ne compte pas. La manière dont sont réalisées les épreuves est fondamentale et en particulier les stratégies et compensations misent en œuvre pour palier à l'incapacité. Nous verrons que la plupart des tests existants ont été dérivés de celui de Caroll.
J'ai sélectionné un certain nombre d'outils et passé en revue leurs caractéristiques en vérifiant s'ils répondaient aux critères exposés au chapitre II. Nous verrons qu'ils sont peu nombreux à répondre à ces critères.

IV.2 Les différents tests.

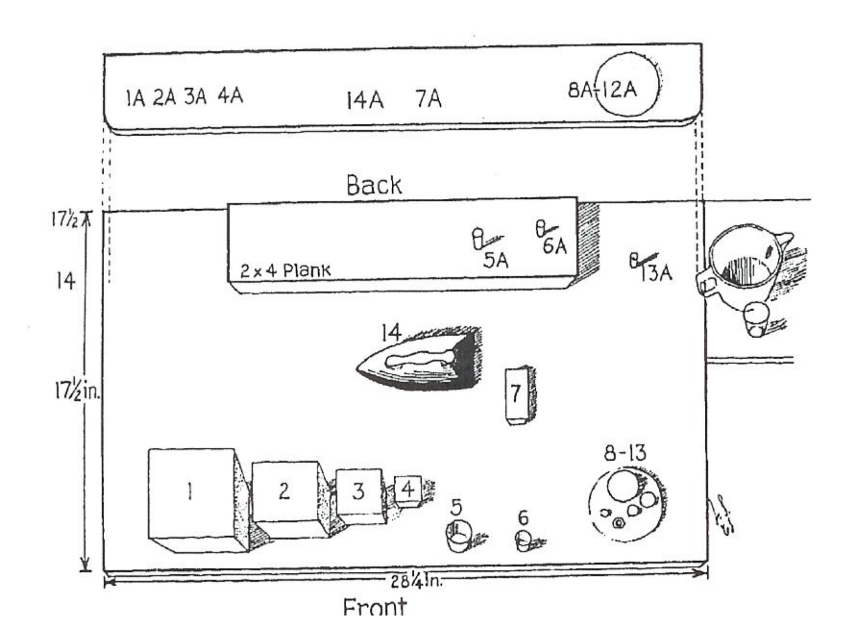
IV.2.1 Le Upper Extremity Function Test (UEFT) de Caroll:

C'est en 1965 que Caroll (19) a décrit ce test à Baltimore. Il était médecin en médecine physique et réhabilitation. Le but du test était de faire réaliser à des patients des activités de la vie quotidienne utilisant la main et le membre supérieur. Ce test devait être facile et rapide à réaliser.
La phase préliminaire du test a consisté à tester deux cent patients variés (post-traumatiques, neurologiques et PR) avec des activités représentant la majeure partie de l'utilisation du membre supérieur dans la vie de tous les jours. Ce pré-test a été comparé avec un test conventionnel du membre supérieur (non précisé dans l'article). La corrélation était très bonne entre les résultats des deux tests. L'auteur ne la chiffre pas, il rapporte la droite de corrélation suivante (cf).



Ce pré score d'après l'auteur donne une bonne indication de l'ensemble de la fonction de la main dans les activités de la vie quotidienne.

Ensuite à partir du pré-test, l'UEFT a été décrit et standardisé.
Ce test consiste à déplacer des objets d'une table sur une étagère (cf schéma), écrire un nom, placer la main à la bouche, derrière le cou et sur la tête, et à verser de l'eau dans un verre. Les objets à déplacer sont de différentes formes et poids et permettent de tester le grasp, le grip, la pince, le placement, l'extension et l'élévation du bras, la prono supination et la force.



Les positions 1, 2,3 et 4 explorent le grasp avec des cubes de différentes tailles et poids (9 à 576 gr).
Les positions 5 et 6 explorent le grip avec des tubes de deux diamètres différents à mettre autour d'un petit piquet de bois.
La position 7 étudie la préhension latérale avec la prise d'un morceau de bois.
Les positions 8 à 12 explorent la pince avec des sphères de différentes tailles. La position 8 est globale et les 9 à 12 testent chaque doigt.
Les positions 13 et14 étudient le placer (rondelle autour d'une pointe) et le déplacement d'un fer à repasser.
Ensuite le patient doit verser de l'eau d'un pichet dans un récipient, puis du récipient dans un autre récipient (mouvement de pronation) puis de nouveau dans le premier récipient (mouvement de supination).
Puis le patient doit placer la main derrière la tête, sur le dessus de la tête et à la bouche.
Enfin il doit écrire son nom.
Ce test représente en tout 33 items.

Pour réaliser ce test le patient doit être assis sur une chaise devant la table et chaque main est testée séparément.
On cote en quatre degrés :
3 : test fait normalement
2 : test complet mais fait lentement
1. test incomplet. Le patient peut déplacer les objets mais n'est pas capable de placer les objets en position correcte. Pour écrire le nom, degré 1 si peut tenir le stylo mais ne peut écrire lisiblement.
0 : le patient ne peut faire aucune partie du test.
Le score maximal est de 99 pour la main dominante et de 96 pour l'autre.

Pour essayer de valider le test, Caroll a fait une étude chez 79 patients, principalement hémiplégiques et en général âgés, représentant 120 mains anormales. Sur ces 79 patients seuls 18 étaient totalement indépendants. Les scores étaient en général très bas dans les trois quarts des cas inférieurs à 89 ce qui signe un score correspondant à une main non ou peu fonctionnelle.
Caroll a testé la fiabilité test-retest et inter observateur mais il ne donne que des chiffres bruts sans analyse statistique. Il signale une fiabilité inter observateur à 97 %.

En conclusion :
On peut dire avec Caroll, que ce test sert à déterminer quelles fonctions sont atteintes et comment elles peuvent changer avec le traitement chirurgical, l'avancée de la maladie et les traitements médicaux. C'est une estimation de la fonction de la main dans les activités de la vie quotidienne qui investigue les principales fonctions (grasp, pince, grip, placer, extension prono supination du bras, flexion/abduction de l'épaule). Bien sûr, ces différents items ne représentent qu'une petite partie des activités d'une main normale dans la vie quotidienne. Les items étudiés sont nécessaires et indispensables pour la plupart des activités de la vie quotidienne. Le test a été conçu pour donner à la fonction préhensile et à la fonction de stabilisation/placement une valeur égale.
Ce test est certainement bien conçu mais le reproche principal à apporter est l'absence d'étude statistique sérieuse en ce qui concerne la fiabilité, la validité de contenu, de construction et la sensibilité au changement. De plus, le choix d'une population de patients âgés, atteints d'affections neurologiques n'a pas été judicieux pour un test censé être plutôt utilisé en pathologie orthopédique et rhumatologique. Des questions restent sans réponses comme celle de savoir à partir de quel chiffre d'amélioration du score total observe-t-on une amélioration significative de la fonction ? L'auteur nous dit que 10 points de plus représentent une amélioration significative mais sans argumenter statistiquement ce chiffre. L'autre problème est qu'il est impossible de dire si deux patients ayant le même score ont le même degré d'incapacité.
A la décharge de Caroll on peut affirmer qu'il y a 40 ans les critères statistiques pour mettre sur pied un outil d'évaluation étaient moins stricts que maintenant.
Ce test a été le point de départ d'autres tests ultérieurs que nous allons voir et je pense qu'à ce titre il était intéressant de s'attarder sur le travail de Caroll.
Le tableau n°5 tiré de l'article de Rudman (77) résume ce test.

PROPRIETES	UEFT(Upper Extremity Function Test)
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	En partie
Unilatéral	Oui
Bilatéral	Non
3) Méthode d'administration	Observation
4) Interprétation des résultats	Vague, analyse des items
5) Spécificité	Populations variées
6) Disponibilité	
Préfabriqué	Non
Domaine public	Oui
Langue de publication	Anglais
Coût	?
7) Temps demandé	
Administration	20-30 minutes
Entraînement du testeur	Minimal
8) Acceptabilité	
Description du but	Non
Approprié à des adultes	Oui
Langue	Anglais
Application dans une autre culture	Large
STANDARDISATION	
1) Instructions	
Administration	En partie
Calcul du score	En partie
Interprétation	En partie
2) Equipement	Non préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif, large étendue
2) Fiabilité	
Inter observateurs	?
Intra observateur	?
Test-retest	?
Cohérence interne	?
3)Validité	
Contenu	?
Construction	?
Relation avec AVQ	oui
Relation avec statut physique	
Distinction des groupes	
Critères	?
Sensibilité au changement	?
4) Normes	non

Tableau n°5 : caractéristiques de l'UEFT (D'après Rudman).

IV.2.2 Le functional hand testing.

Greenseid (37), chirurgien plasticien, décrit en 1968 douze tests pour évaluer la fonction active de la main. Ces tests évaluent la pince, le grip , le grasp et la manipulation digitale. Certains tests sont chronométrés et d'autres sont qualitatifs et sont notés bon, suffisant ou échec.
Les 12 tests sont les suivants :

- Ramasser des vis et les transférer dans une coupelle.
- Ouvrir une serrure à codes.
- Dessiner.
- Faire un nœud.
- Faire des lacets de chaussures.
- Couper avec une paire de ciseaux un modèle.
- Tenir un verre et le remplir d'eau avec une cruche (2 tests).
- Couper un fil de fer.
- Ouvrir une porte.
- Ouvrir des récipients.
- Attraper une balle.

On le voit, ces activités sont proches des activités de la vie quotidienne mais malheureusement l'auteur ne l'a testé que chez 14 patients et aucune analyse statistique ne vient valider ce test. De plus, les compensations et astuces mises en œuvre par le patient ne sont pas prises en compte.
Ce test ne peut être retenu dans notre analyse mais de nombreux items décrits dans ce test ont été repris ultérieurement.

IV.2.3 Le Jebsen Hand Function Test (JHFT).

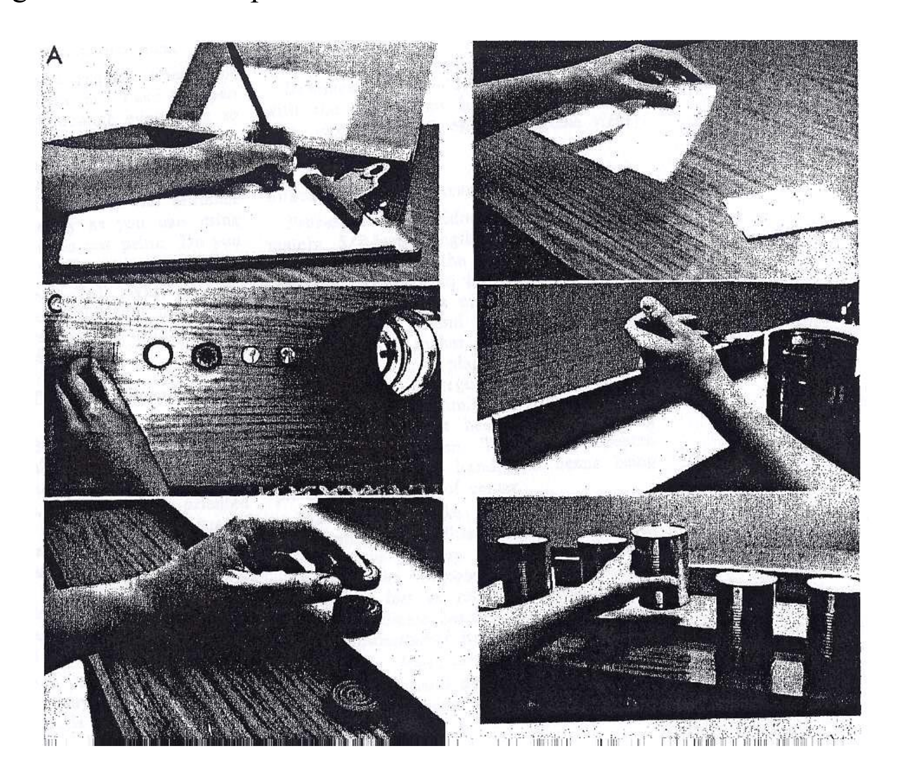
En 1969, Jebsen, médecin de médecine physique, décrit ce test objectif et standardisé (46).
Les buts de ce test étaient :

- De donner des mesures objectives de tâches standardisées que l'on puisse comparer à des normes.
- D'estimer la plupart des aspects de la fonction de la main utilisée dans des activités de la vie quotidienne.
- De suivre la capacité fonctionnelle dans chaque catégorie de main atteinte, par des tests répétés.
- D'être administré rapidement.
- D'utiliser un matériel simple et disponible facilement.

Méthodes :
Sept tests ont été choisis pour couvrir la fonction de la main. L'auteur ne signale pas comment s'est opéré le choix.
Les sept tests choisis sont les suivants. A noter qu'ils sont tous chronométrés et que l'on teste la main non dominante puis la main dominante. Les tests sont donc tous unilatéraux:

1. Ecrire une phrase de 24 lettres en lettres majuscules.
2. Tourner des cartes (simulation de l'action de tourner les pages d'un livre).
3. Manipulation de petits objets : boîte de conserve vide, trombones, pièces de monnaie. L'épreuve consiste à mettre les pièces et les trombones dans la boîte de conserve.

4. Simulation d'alimentation : le patient doit prendre avec une petite cuillère 5 haricots secs et les mettre dans une boîte de conserve vide.
5. Empiler des jetons : 4 gros jetons en bois sont mis côte à côte sur la table et doivent être empilés.
6. Manipulation de gros objets légers : 5 boîtes de conserves vides de grandes tailles sont à déplacer de 15 cm environ.
7. Manipulation de gros objets lourds : 5 boîtes de conserve pleines de grandes tailles de 500 grammes sont à déplacer de 15 cm environ.



Pour établir des normes, 300 sujets répartis en 30 femmes et 30 hommes dans les tranches d'âge suivantes : 20-29 ans, 30-39 ans, 40-49 ans, 50-59 ans et 60-94 ans ont été testés. La moyenne et l'écart type ont été calculés avec analyse de la variance pour l'âge et le sexe.
Pour évaluer la fiabilité du test, 26 patients ayant une atteinte stable de la main, ont été testés en deux occasions. Les résultats ont été analysés en obtenant le coefficient de corrélation de Pearson.
Pour calculer l'effet d'apprentissage, les résultats du test-retest obtenus ont été évalués sous l'angle de l'amélioration de la performance lors du deuxième test, en calculant le t-test de la moyenne des différences et le Chi2 du pourcentage de sujets s'améliorant.
Pour étudier la capacité discriminative du test, 33 patients (12 hémiplégiques, 10 PR et 11 traumatismes cervicaux) ont été testés et comparés aux sujets normaux.

<u>Résultats :</u>

Sujets normaux : il existe des différences significatives pour l'âge et le sexe. Les sujets âgés réalisent les tests plus lentement que les sujets jeunes, et les femmes sont plus rapides que les hommes. Deux groupes d'âge ont été faits pour l'analyse : 20 à 59 ans et 60 à 94 ans. La différence entre les moyennes par groupes d'âge, par sous-tests et mains était statistiquement significative (t-test) avec $p < 0.005$ pour tous les cas sauf pour la simulation d'alimentation (test n° 4) pour les femmes. Il faut noter que les temps réalisés sont tous inférieurs à 10

secondes sauf pour écrire la phrase. Il en découle donc que les écarts types sont aussi petits. Les scores de 90% des sujets normaux tombent à l'intérieur de deux écarts types de la moyenne sauf pour les objets grands et légers chez les sujets de 60 à 94 ans.

Patients : la fiabilité test-retest appréciée par le coefficient de corrélation de Pearson est bonne. Les coefficients vont de 0.67 à 0. 99 pour la main dominante, et de 0.60 à 0.92 pour la main non dominante ($p<0.001$).
L'effet d'apprentissage apprécié par le t-test et le Chi2 chez les 26 patients a montré qu'il n'y avait pas d'effet d'apprentissage.
Le test discrimine bien les sujets normaux des patients.

Discussion :

D'après Jebsen, ce test est une évaluation objective et standardisée des principaux aspects de la fonction de la main. Mais ce test souffre de nombreuses imperfections statistiques que Rudman (77) a bien souligné et qui sont résumées dans le tableau 6 ci-après.
Il faut aussi souligner que seul le temps de réalisation du test est pris en compte et à aucun moment les compensations ne sont évaluées et notées.

D'autre études (40,79, 85) ont été faites avec le test de Jebsen.
Dans les années 1980, plusieurs études (citées in 85), ont appliqué le Jebsen test dans des situations cliniques variées : immobilisation du poignet, paralysie cérébrale de l'enfant, hémiplégie, synostoses radio ulnaires, arthrose de la main et orthèses d'extension du poignet.
Hackel (40) en 1992, a calculé de nouvelles normes pour les sujets âgés en définissant de nouvelles catégories d'âge : 60-69 ans, 70-79 ans et 80-89 ans. Sur 121 patients il a été mis en évidence que la fonction de la main diminuait avec l'âge que ce soit chez l'homme ou chez la femme. La fiabilité inter observateur a dans cette étude été calculée par le coefficient de corrélation intra classe. Elle était de 0.82 à 1.00. La fiabilité intra observateur était de 0.85 (coefficient de corrélation de Pearson).
Stern (85) en 1992, a étudié la stabilité de la mesure sur trois sessions. D'après elle, cette fiabilité test retest est très bonne, mais l'analyse statistique n'est pas détaillée et cet article est pour le moins flou.
Sharma (79) en 1994, a utilisé le test dans la PR. 25 patients ont été testés. Le JHFT détecte bien les différences entre les patients malades et les sujets sains. Pour les items du JHFT, ils montrent tous, sauf l'écriture, une corrélation avec les AVQ et les déformations dues à la maladie. Par contre, il n'y a pas de corrélation avec la douleur.
Nous avons aussi vu dans le chapitre précédent, que le JHFT avait été choisi comme gold standard et comparé au NKHDT et au FDT.

En conclusion :
Le JHFT, même si ses intentions étaient au départ bonnes, ne montre pas des qualités suffisantes pour être utilisé sans réserve et surtout comme « gold standard ». Nous avons vu que ses propriétés psychométriques ont été peu évaluées et sont insuffisantes, ce qui n'a pas empêché que ce test soit utilisé dans beaucoup d'études, surtout dans les pays anglo-saxons, et considéré comme une référence. A mon sens le défaut principal est aussi que l'unité de mesure soit le temps et qu'aucune directive ne soit donnée pour une appréciation qualitative des tests. Je pense qu'en l'état, ce test devrait être abandonné avant que des études s'intéressent de près à démontrer qu'il a toutes les qualités statistiques requises.

PROPRIETES	JHFT (Jebsen Hand Function Test)
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	En partie
Unilatéral	Oui
Bilatéral	Non
3) Méthode d'administration	Observation
4) Interprétation des résultats	Normes
5) Spécificité	Populations variées
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	Anglais
Coût	260 $
7) Temps demandé	
Administration	15 minutes
Entraînement du testeur	Minimal
8) Acceptabilité	
Description du but	Non
Approprié à des adultes	Oui
Langue	Anglais
Application dans une autre culture	Quelques limitations
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	Oui
2) Equipement	Préfabriqué
BUTS	Descriptif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif, large champ
2) Fiabilité	
Inter observateurs	Limitée
Intra observateur	?
Test-retest	Limitée pour la majorité des sous-tests
Cohérence interne	?
3) Validité	
Contenu	?
Construction	
Relation avec AVQ	Oui
Relation avec statut physique	Inconsistante (dans certaines études oui dans d'autres non)
Distinction des groupes	Limitée
Critères	?
Sensibilité au changement	Limitée
4) Normes	Oui discutables pour les sujets de moins de 60 ans, acceptables pour les + de 60 ans.

Tableau n°6 : caractéristiques du JHFT (D'après Rudman).

IV.2.4 Le Smith hand function evaluation (SHFE).

Smith (81) en 1973 a développé un nouveau test original par sa structure. Elle exerce comme ergothérapeute à l'université du Michigan. Le SHFE est une évaluation standardisée de la coordination de la main et des activités fonctionnelles de la main associées aux AVQ. Il donne une évaluation de la force et de la dextérité uni et bi manuelle. C'est un test à la fois descriptif et évaluatif, puisqu'il peut être ré administré pour suivre l'évolution après une thérapie. L'auteur signale que ce test ne permet pas de mesurer objectivement : la coordination motrice œil main, les mobilités de l'épaule ou du coude, les types de grasp utilisés, le déficit visuel, la fatigabilité ou le temps de réaction. Par contre, le thérapeute est invité à observer cela et à le noter. Le temps d'administration est de 20 minutes.

Matériels et méthodes :

91 personnes saines, 51 femmes et 40 hommes de 20 à 60 ans. Le sujet est assis. Les activités sont chronométrées en secondes.
Quatre sections dans le test :

- Section A : Tâches uni manuelles avec la main dominante puis l'autre main.

1) Déplacer des cubes de différentes tailles soit à droite soit à gauche en fonction de la main testée.
2) Mettre 4 clous de tailles différentes dans un verre.
3) Prendre quatre pièces de monnaie de tailles différentes et les mettre dans un verre.
4) Mettre 8 grosses chevilles et 8 petites sur un tableau à trous.

- Section B : Activités de la vie quotidienne. Utilisation des deux mains. Ces activités sont présentées sur un tableau.

1) Epingle à nourrice : la prendre, l'ouvrir puis la fermer.
2) Ceinture : la déboucler et la reboucler.
3) Boutons : déboutonner et reboutonner 3 boutons de plastique de différentes tailles.
4) Fermeture éclair : ouvrir et fermer.
5) Nœud : faire un double nœud.
6) Faire un nœud de papillon (avec boucles).
7) Lacer des chaussures en passant dans les trous.

- Section C : Ecrire. Le sujet fait sa signature et trace un rectangle puis une ligne courbe.

- Section D : Force au dynamomètre en Kg. Deux essais de chaque côté.

Résultats :

Aucune analyse statistique n'est rapportée en dehors des résultats avec les moyennes et les écarts types des temps mis pour les différents tests. Les trois hypothèses de départ sont confirmées : 1) les droitiers sont plus habiles que les gauchers (mais il n'y avait que 9 gauchers dans cette étude) ; 2) les hommes et les jeunes adultes ont plus de force que les plus âgés ; 3) les femmes et les jeunes adultes démontrent une meilleure dextérité et travaillent

plus vite que les hommes et les adultes âgés. Smith rapporte aussi que le chronomètre n'est pas un instrument exact pour mesurer la performance, car ce test étant court, le temps pour presser le bouton et donner simultanément le départ peut ajouter des secondes au test et fausser le résultat.

Il n'est pas retrouvé de trace dans la littérature du test de Smith sauf en 1993 ou un article de Jarus (45) a étudié le JHFT et le SHFE. Cet auteur pense qu'en combinant les deux tests, une meilleure évaluation de la main serait possible.

<u>*Méthodes :*</u>

144 volontaires sains (70 hommes et 74 femmes) et 21 volontaires (3 hommes et 18 femmes) avec fracture de Pouteau-Colles.
Les tests ont été traduits en hébreu et administrés selon le protocole de la publication princeps. Les deux tests combinés représentent 34 sous-tests.
Le temps pour faire les deux tests a été de une heure environ.
L'analyse des résultats a été faite en trois stades. Le premier (l'analyse de facteur) a consisté à réduire les 34 sous-tests à un petit nombre puis de les grouper entre items hautement corrélés et à les appeler facteurs. Le deuxième a consisté à déterminer la fiabilité de chaque facteur et d'exclure les sous-tests non fiables. Le troisième a consisté à tester la capacité des tests à prédire la fonction de la main en comparant la moyenne du groupe sans fracture à celui du groupe avec fracture (utilisation du t-test).

<u>*Résultats :*</u>

L'analyse de facteur indique que quatre facteurs rendent compte de 5% de la variance ce qui est bien. Ces quatre facteurs sont : le pinch, le grasp, la précision sur la cible et les AVQ. 18 sous-tests ont été exclus. La fiabilité de chaque facteur allait de 0.8 à 0. 96. Le facteur 1 comprend 4 items appartenant au SHFE (chevilles), le facteur 2 comprend 4 items appartenant au JHFT (boites de conserves vides et pleines), le facteur 3 comprends 6 items 4 appartenant au SHFE (cubes et clous) et 2 appartenant au JHFT (pions en bois), enfin le facteur 4 comprend 2 items appartenant au JHFT (ramasser les haricots). Il a été calculé la moyenne de chaque facteur dans chaque groupe de patients et une analyse multi variée a été pratiquée (MANOVA). Celle-ci révèle une différence significative entre les deux groupes de patients avec et sans fracture. Ceci veut donc dire que les quatre facteurs peuvent distinguer une population de l'autre.

Les auteurs pensent qu'avec la combinaison des deux tests, du moins des items les plus relevants identifiés en quatre facteurs, on a à disposition un test plus performant que le JHFT ou le SHFE seuls. Mais, tous les types de préhensions ne sont pas réunis dans ce test. De plus, il n'évalue pas la qualité du mouvement et l'aspect statique de la préhension. Enfin, ce nouveau test n'a pas été étudié en ce qui concerne la validité, la fiabilité test retest, la fiabilité inter observateurs.

En conclusion :

Le SHFE est intéressant dans le concept, et nous verrons plus loin qu'il a été repris, mais il souffre d'une mise au point qui n'a pas répondu aux critères indispensables pour en faire une évaluation de qualité. Le tableau ci-dessous résume cela.

PROPRIETES	SHFE (Smith Hand Function Test)
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	En partie
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Observation
4) Interprétation des résultats	Normes
5) Spécificité	Populations variées
6) Disponibilité	
Préfabriqué	Non
Domaine public	Oui
Langue de publication	Anglais
Coût	?
7) Temps demandé	
Administration	20-60 minutes
Entraînement du testeur	Minimal
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	Anglais
Application dans une autre culture	Quelques limitations
STANDARDISATION	
1) Instructions	
Administration	En partie
Calcul du score	Oui
Interprétation	Oui
2) Equipement	Non préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Non spécifié
2) Fiabilité	
Inter observateurs	?
Intra observateur	?
Test-retest	?
Cohérence interne	?
3) Validité	
Contenu	?
Construction	?
Relation avec AVQ	
Relation avec statut physique	
Distinction des groupes	
Critères	?
Sensibilité au changement	?
4) Normes	Oui

Tableau n°7 : caractéristiques du SHFE (D'après Rudman).

IV.2.5 Physical capacities evaluation of hand skill (PCE).

En 1976, Bell (12) a publié ce test qu'elle utilisait déjà depuis des années. Il s'agit d'un test utilisé en neurologie, qui mesure la dextérité de la main. C'est un test composite avec plusieurs sous tests, 5 tests unilatéraux et 7 tests bilatéraux. Des normes ont été déterminées pour des gens normaux et pour trois sous populations de patients neurologiques : paraplégie, hémiplégie droite, hémiplégie gauche.
Je n'ai pas retenu ce test car aucune analyse statistique n'est disponible dans cette publication.

IV.2.6 L'Action Research Armtest (ARA).

Publié par Lyle en 1981 (59), ce test est surtout utilisé en neurologie. Il s'agit d'un reformatage du test UEFT de Caroll décrit plus haut. En effet, l'auteur estime que l'UEFT est trop long à faire, environ une heure pour les deux membres supérieurs, et que de nombreux items mesurent en fait plus de choses que ce qu'ils étaient censés mesurer lors de la description princeps du test. De plus, l'auteur estime que de trop nombreux items évaluent la dextérité fine opposition du pouce- doigts.

Matériel et méthodes :

20 patients (13 hommes et 7 femmes), tous hémiplégiques depuis une moyenne de 46 mois, âgés de 26 à 72 ans.
Les patients ont passé le test UEFT sauf l'item 33 (écriture). Les patients ont été testés deux fois par deux équipes indépendantes de deux personnes avec un intervalle de test de 7,5 jours. Chaque main est testée.

Résultats :

La fiabilité test retest est supérieure à 0.98 dans la plupart des cas. Si on calcule séparément ces corrélations en fonction des sous-tests, les coefficients de corrélation de Pearson vont de 0.89 à 0.99 pour la main atteinte.

A partir de là, l'auteur explique comment il a réduit l'UEFT pour en faire l'ARA. Pour ce faire, il utilise un outil statistique la « Guttman scale analysis » ou analyse d'échelle de Guttman qui permet de réduire le nombre d'items. Puis, les items sont regroupés en fonction de leur corrélation. Dans l'UEFT certains items n'ont pas une grande cohérence interne inter item (pincer, placer, prono supination) et d'autres items ne sont pas placés dans la bonne sous échelle. L'analyse statistique corrélative des items a permis de dégager quatre sous échelles concernant l'action de saisir, d'agripper, de pincer et la motricité grossière. Pour chaque échelle, les items ont été classés hiérarchiquement en fonction de leur difficulté. On peut donc prédire la réussite à un item moins difficile à partir de la réussite à un item le plus difficile ; à l'inverse l'échec à un item moins difficile prédit l'échec à un item plus difficile.
Toutes les sous échelles répondaient aux critères statistiques de l'analyse d'échelle de Guttman avec un coefficient de reproductibilité supérieur à 0.9. Une simulation sur ordinateur a ensuite permis de procéder à une réduction de la longueur du test allant de 67 % à 87 %. Les sous échelles ont été corrélées entre elles et avec le score total de l'ARA. Les coefficients de corrélation de Pearson étaient tous significatifs avec un $p<0.001$ et des valeurs de 0.7 à 0.95. Une épreuve de validation croisée de l'ARA a confirmé que cette procédure d'évaluation

hiérarchique était pertinente en reprenant (analyse statistique seule) les 50 patients initiaux de Caroll avec des diagnostics plus diversifiés (amputations, PR..).
D'après l'auteur, l'ARA est un bon test pour la réadaptation de l'hémiplégie. Il permet de mesurer les progrès obtenus grâce aux traitements, d'évaluer un traitement soit en clinique soit en recherche.
Il faut noter que la validité de contenu, de construction et la sensibilité au changement n'ont pas été étudiés dans cette publication, que l'ARA est dérivé de l'UEFT dont on a dit plus haut qu'il n'avait pas été étudié de manière statistiquement bonne et que l'étude initiale de l'ARA n'a été faite que chez 20 patients.
Néanmoins, ce test s'est imposé dans les milieux neurologiques et est devenu le test de référence dans de nombreuses publications ces 20 dernières années en ce qui concerne la neurologie centrale (40 publications recensées environ). Je citerais deux publications intéressantes sur l'ARA.
La première de 2001 a étudié la fiabilité intra et inter observateur sur 20 patients neurologiques (88). Cette fiabilité était excellente supérieure à 0.98. Dans cette étude était aussi confirmé que l'ARA avait la capacité de détecter une différence cliniquement pertinente de 5,7 points.
La deuxième (89) de 2002 a montré que l'ARA est une échelle unidimensionnelle au contraire de ce que Lyle avait cherché à montrer en créant quatre sous échelles ! Cet auteur propose de réduire le test à 15 items au lieu de 19.

En conclusion :
L'ARA a été mieux étudié statistiquement que l'UEFT dont il est une version courte, mais de nombreux paramètres statistiques importants n'ont pas été décrits en particulier la validité et la sensibilité au changement. Malgré cela, ce test s'est imposé dans les milieux de rééducation neurologique surtout en raison de sa simplicité et de sa rapidité d'évaluation. Il faudrait des études complémentaires sur les caractéristiques statistiques pour que l'on puisse l'utiliser sans doute. Il faut noter que ce test n'a été utilisé qu'en neurologie ce qui réduit encore son intérêt et que si son utilisation devait se faire dans d'autres pathologies, il faudrait refaire d'autres études pour le valider dans cette indication. Ci-dessous le mode d'emploi de l'ARA et un tableau résumant ses caractéristiques.

Mode d'emploi de l'ARA :
Le score est de 3 si l'activité demandée est réalisée normalement, de 2 si l'activité est faite mais avec grande difficulté ou trop lentement, de 1 si l'activité est faite partiellement et de 0 si elle n'est pas faite.
Chaque main est testée séparément. Le test n'est pas chronométré.

1) Sous test SAISIR (GRASP) :

- cube de 10 cm (le plus difficile) : si score 3 passer directement au sous-test 2 et donner 18 (6x3) de score total au 1.
- Cube de 2,5 cm (le plus facile): si score 0 le total sera 0 et passer au sous test 2.
- Cube de 5 cm.
- Cube de 7,5 cm.
- Balle de 7,5 cm.
- Pierre d'aiguisage 10 cm x 2,5 cm x 1 cm.

2) Sous test AGRIPPER (GRIP) :

- Verser de l’eau d’un verre dans un autre verre (pronation) : si score de 3 le score total sera 12 et passer au sous test 3.
- Tube de 2,25 cm : si score 0 le score total sera 0 et passer au sous test 3.
- Tube de 1 cm.
- Mettre une rondelle autour d’une cheville.

3) Sous test PINCER (PINCH) :

- Bille de 6 mm entre pouce et annulaire : si score de 3 le score total sera de 18 et passer à sous test 4.
- Bille de marbre entre pouce et index : si score de 0 le score total est de 0 et passer au sous test 4.
- Bille de 6 mm entre pouce et majeur.
- Bille de 6 mm entre pouce et index.
- Bille de marbre entre pouce et annulaire.
- Bille de marbre entre pouce et index.

5) Sous test MOTRICITE GROSSIERE :

- Placer la main derrière la tête : si score de 3 le score total sera de 9 et si score de 0 le score total sera de 0.
- Placer la main sur la tête.
- Placer la main à la bouche.

Le score maximal est de 57. Plus le score est bas plus le résultat est mauvais.

PROPRIETES	ARA (Action Research Arm test)
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	En partie
Unilatéral	Oui
Bilatéral	Non
3) Méthode d'administration	Observation
4) Interprétation des résultats	Vague
5) Spécificité	Hémiplégie
6) Disponibilité	
Préfabriqué	Non
Domaine public	Oui
Langue de publication	Anglais
Coût	?
7) Temps demandé	
Administration	Court
Entraînement du testeur	Minimal
8) Acceptabilité	
Description du but	Non
Approprié à des adultes	Oui
Langue	Anglais
Application dans une autre culture	Oui
STANDARDISATION	
1) Instructions	
Administration	En partie
Calcul du score	Oui
Interprétation	Non
2) Equipement	Non préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	> 0.98
Intra observateur	>0.98
Test-retest	0.89 à 0.99 main dominante
Cohérence interne	Pas étudiée
3) Validité	
Contenu	?
Construction	?
Relation avec AVQ	
Relation avec statut physique	
Distinction des groupes	
Critères	?
Sensibilité au changement	?
4) Normes	Pas besoin

Tableau n°8 : caractéristiques de l'ARA.

IV.2.7 Le bilan 400 points.

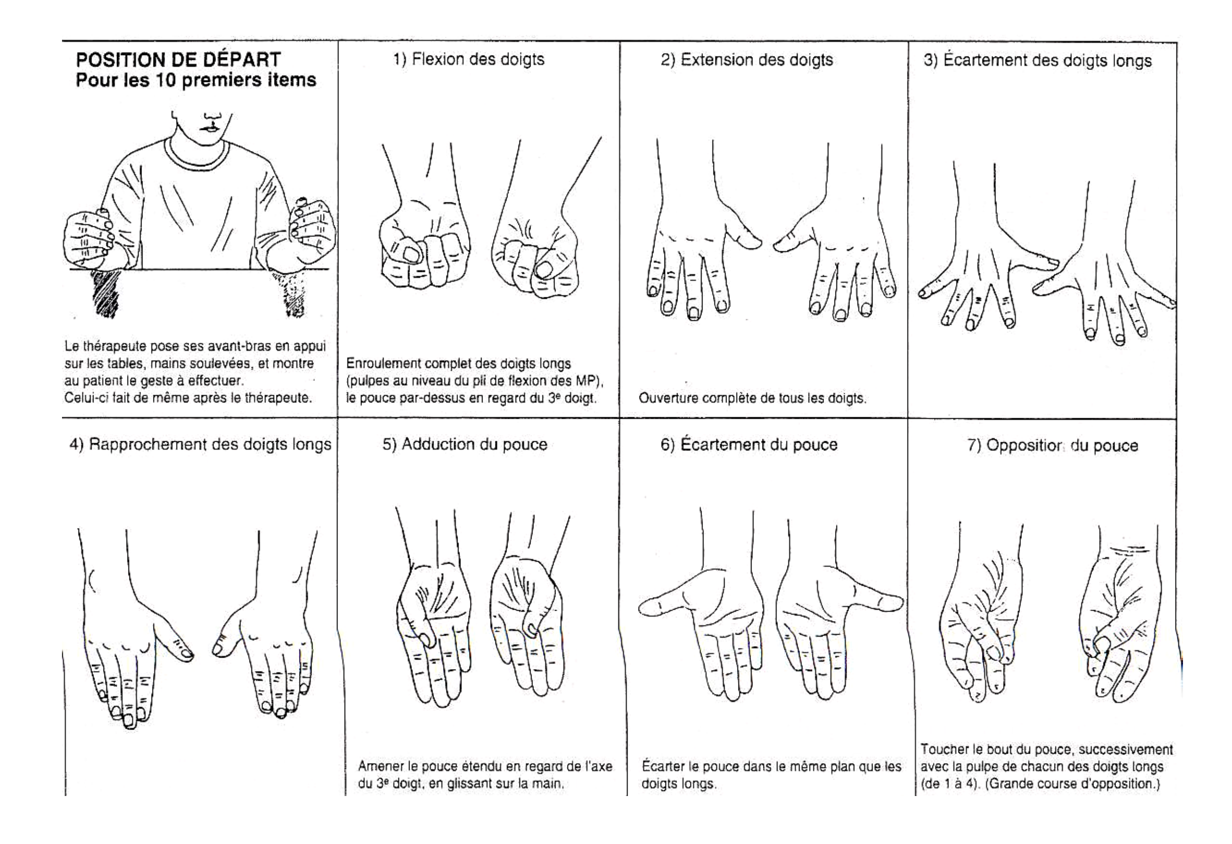
Ce bilan a été conçu par Colette Gable ergothérapeute et le Dr Jacques Xenard membres de l'équipe du centre de rééducation fonctionnelle secteur main de Nancy en 1985. Il a été publié une première fois en 1989 dans une revue de réadaptation professionnelle française. Ensuite en 1993 et 1996 (31,58) deux études de validation ont été publiées. En 1996 et 1997 (32,33) la publication du bilan 400 points sous sa forme actuelle a été faite. Il y a donc eu 11 ans entre la mise au point du bilan et sa publication sous sa forme définitive.
Il est inspiré de différents bilans américains (37, 42, 46) en particulier le Jebsen test (JHFT) et le hand functional testing de Greenseid que nous avons déjà vu dans ce même chapitre.

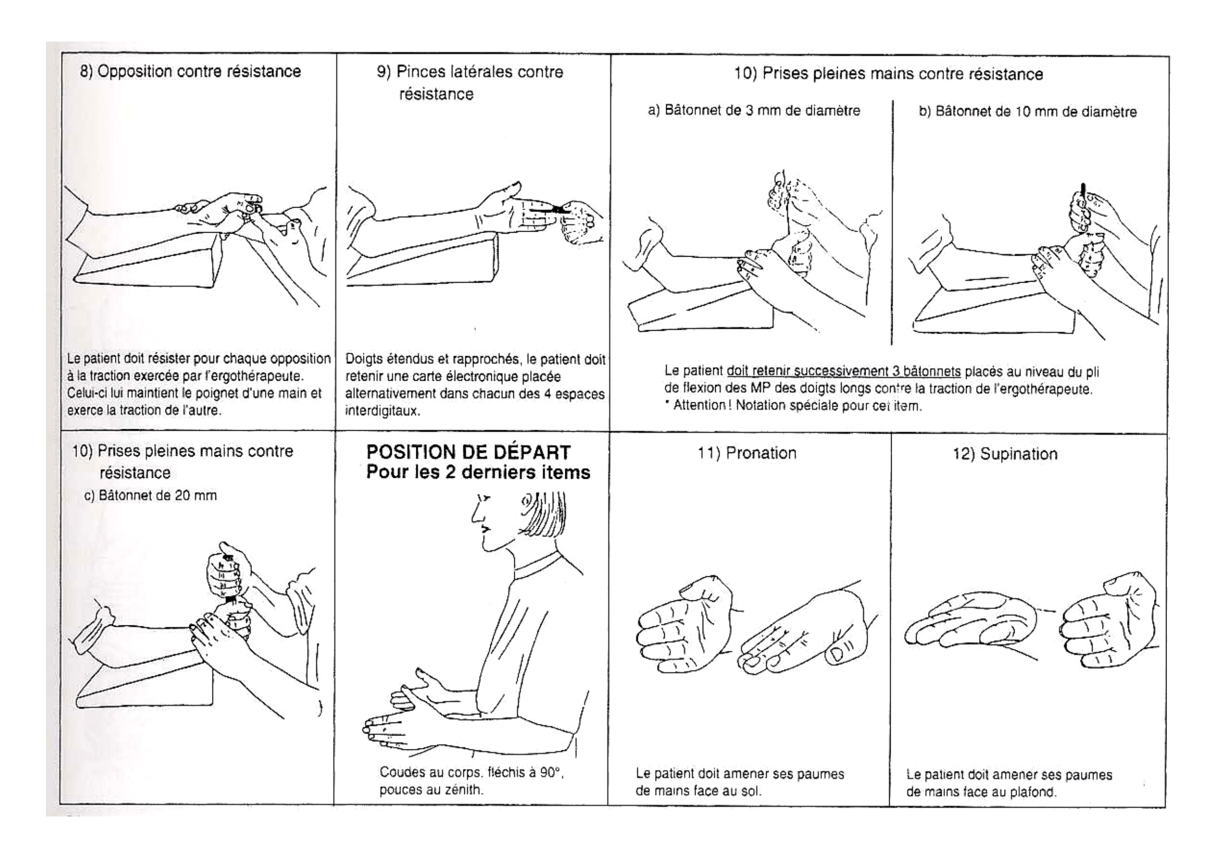
IV.2.7.1 Description du test :

La description de ce test est faite dans l'article du journal d'ergothérapie de 1996 (32). Il est parfaitement standardisé avec un protocole précis.
Il se compose de quatre épreuves chacune notée sur 100 points. Pour les épreuves 1 2 et 4 il faut une table, une chaise pour le patient et un tabouret pour le thérapeute, qui doivent être réglables en hauteur.
Les épreuves 1 et 2 ne sont pas des épreuves comportant des activités de la vie quotidienne mais étudient le mouvement global et la force.

EPREUVE 1 : MOBILITE DE LA MAIN.

On teste simultanément les 2 mains du patient au moyen de 12 mouvements globaux dont on note la qualité d'exécution (cf. schémas). Les schémas sont tirés de la référence 96.





La cotation se fait de 0 à 3.
Note 3 : mouvement normal, conforme au modèle.
Note 2 : mouvement normal, effectué en entier mais lent ou disharmonieux.
Note 1 : mouvement non conforme car effectué de façon incomplète ou avec exclusion ou avec compensation.
Note 0 : mouvement impossible à réaliser soit parce qu'il y a une contre indication soit parce qu'il y a une incapacité provisoire ou définitive.
Des coefficients modérateurs sont attribués aux mouvements car ils sont d'importance inégale. (cf.tableau)

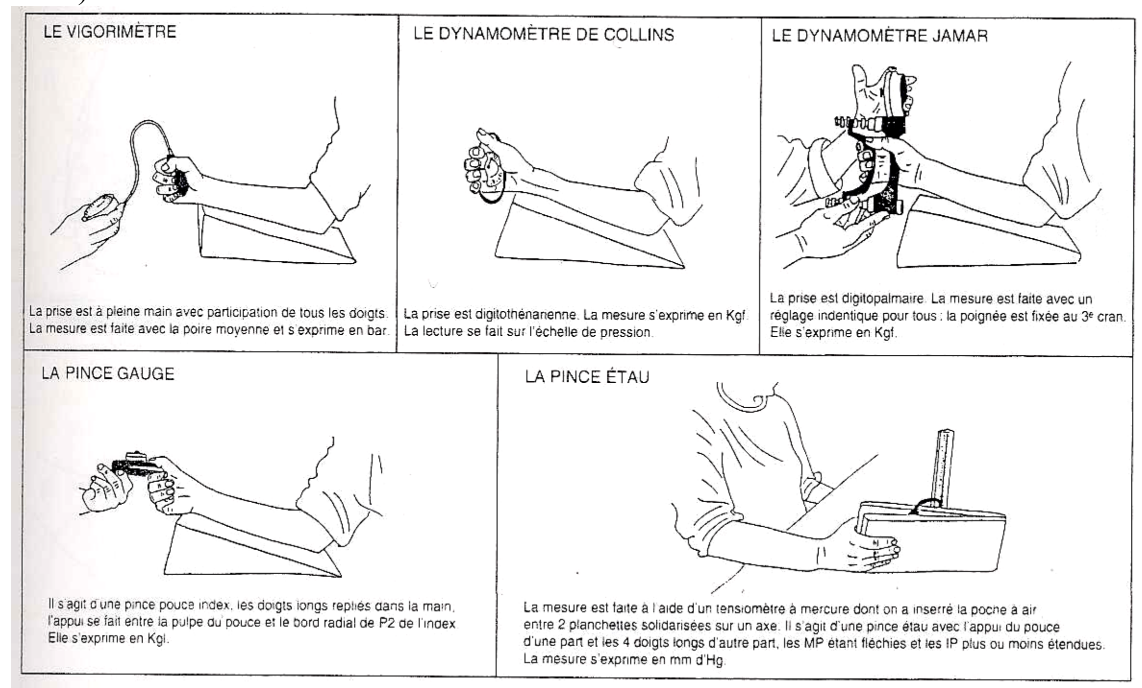
Tableau I. – *Coefficients modérateurs.*

– Flexion des doigts	3	– Opposition du pouce	3
– Extension des doigts	2	– Opposition contre résistance	2
– Ecartement des doigts longs	3	– Pinces latérales contre résistance	2
– Rapprochement des doigts longs	2	– Prises pleine main contre R	3
– Adduction du pouce	2	– Pronation	3
– Ecartement du pouce	3	– Supination	2

Pour le calcul, chaque cotation est multipliée par son coefficient puis on additionne les 12 notes ce qui correspond à la somme des notes côté lésé. Le côté sain a par définition une note maximale de 90 points. Enfin la note est divisée par 90 et multipliée par 100 pour aboutir au pourcentage de mobilité de la main lésée par rapport à la mobilité d'une main saine.

EPREUVE 2 : FORCE DE PREHENSION

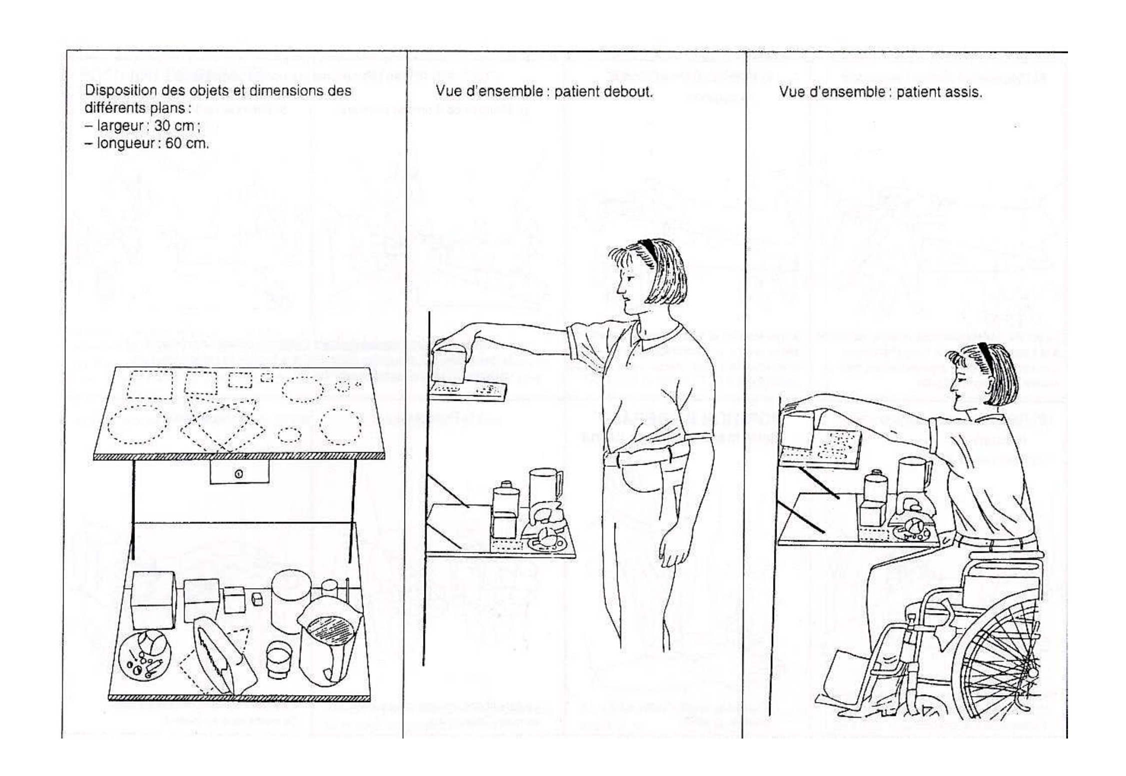
Elle est réalisée au moyen de 5 appareils de mesure du côté sain puis du côté lésé (cf. schémas).

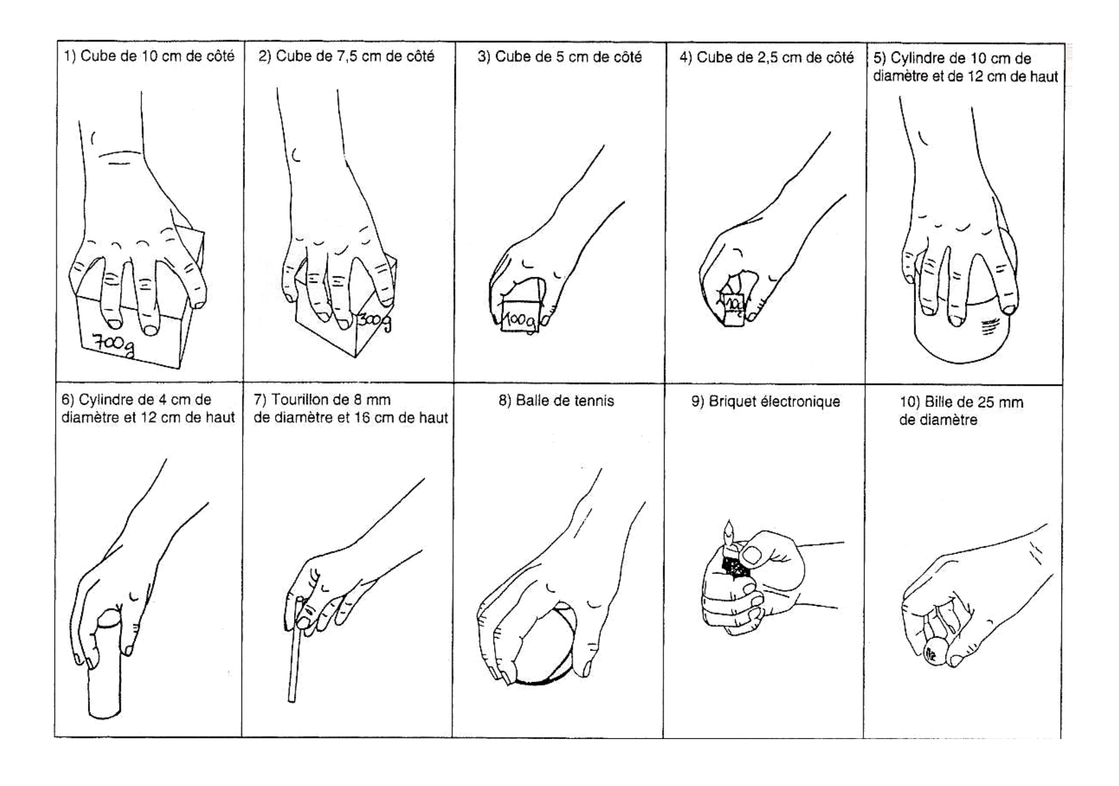


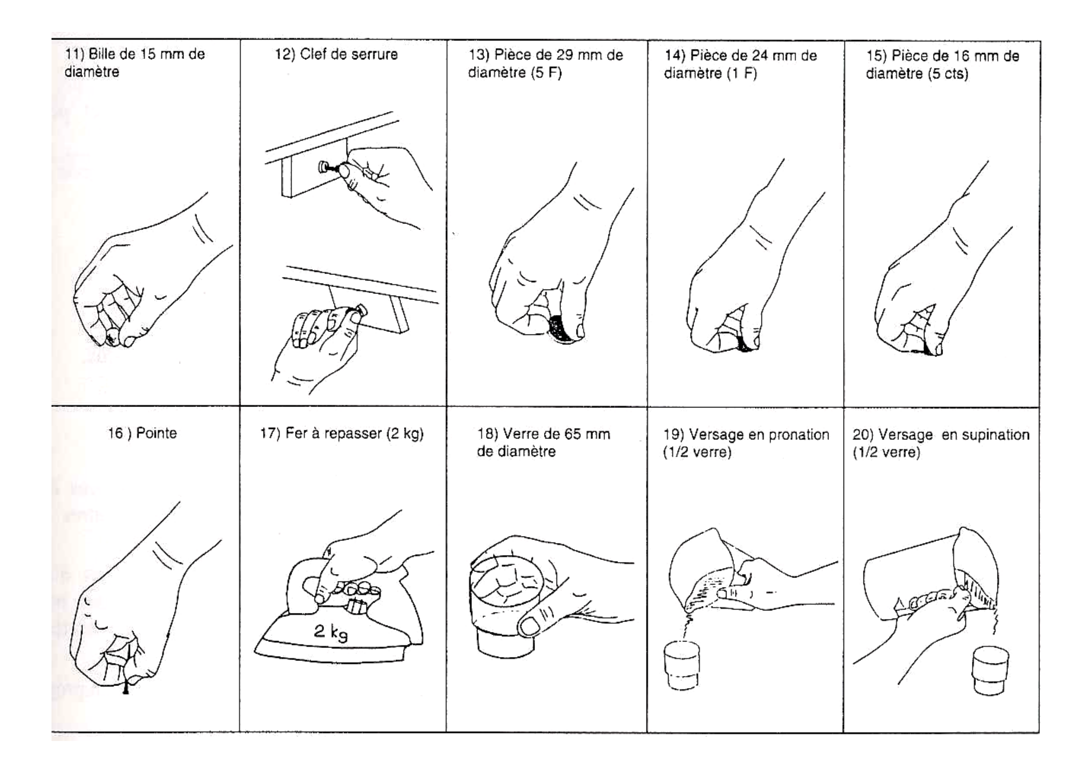
Le patient est installé avec les épaules en adduction rotation neutre, l'avant bras en appui sur le coussin triangulaire ou sur la table pour la pince étau, le coude à 90° et la main à l'extérieur du coussin sans appui.
Le score est déterminé par la moyenne de trois mesures successives.
Pour le calcul du score, la main saine a une force maximale soit 10 points. La main lésée est notée par rapport à la main saine en faisant une règle de 3 : note côté lésé= valeur côté lésé : valeur côté sain X 10.
En règle générale, la main dominante a 10% de force de plus que la main d'appoint. Si la main dominante est lésée il faut retrancher 10% au résultat obtenu. Si la main d'appoint est lésée il faut ajouter 10 % au résultat obtenu. Si les deux mains sont lésées il faut comparer les résultats obtenus avec le groupe témoin de sujets sains (30 hommes et 30 femmes).
Il faut appliquer un coefficient 2 aux résultats pour arriver à une note sur 100.

EPREUVE 3 : PRISE MONOMANUELLE ET DEPLACEMENT D'OBJETS.

Cette épreuve teste la capacité du sujet à prendre 20 objets différents posés sur un plan de référence, dans un ordre précis et à les transporter sur un plan en hauteur.
L'épreuve est chronométrée avec d'abord la main saine puis la main lésée. L'ergothérapeute explique l'épreuve au patient. Cette épreuve doit être faite en moins d'une minute. Le fait que l'épreuve soit chronométrée oblige le patient à faire des gestes spontanés. Le thérapeute doit noter les défauts de préhension du patient.
L'épreuve se fait en position debout ou assise (fauteuil roulant) avec un plan réglable en hauteur. Le matériel et les tests sont représentés sur les schémas ci-dessous.


Disposition des objets et dimensions des différents plans :
– largeur : 30 cm ;
– longueur : 60 cm.
Vue d'ensemble : patient debout.
Vue d'ensemble : patient assis.

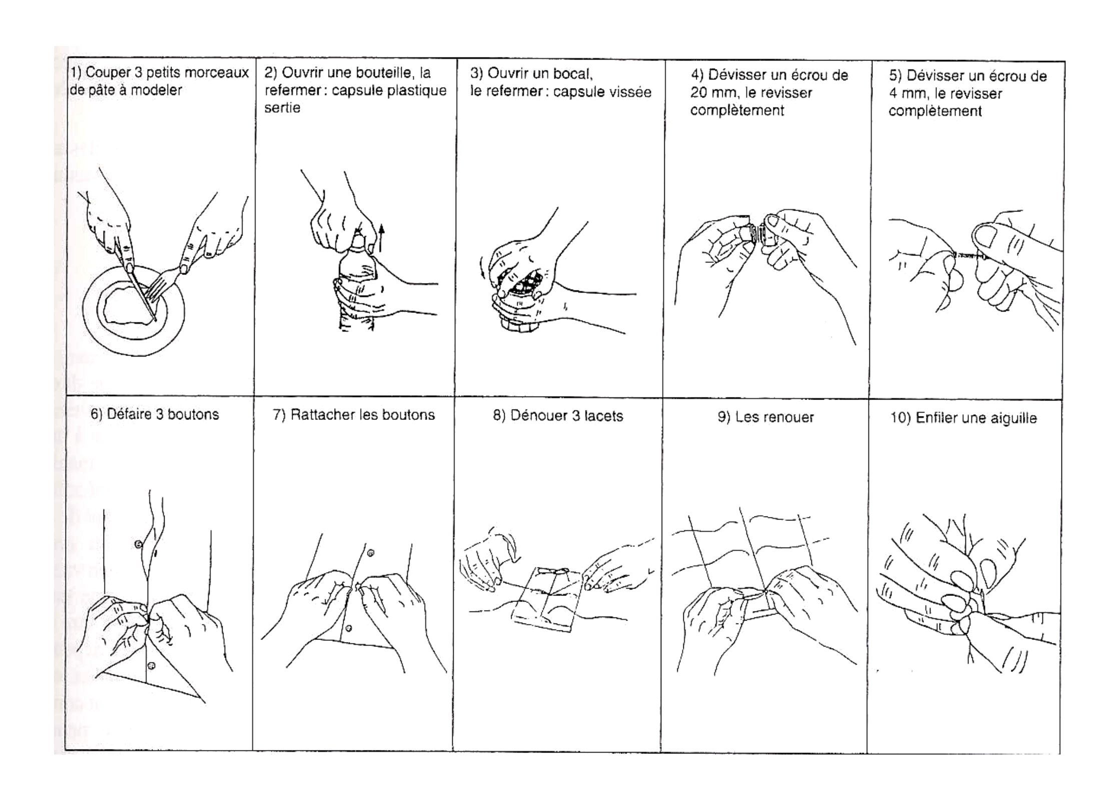

1) Cube de 10 cm de côté
700g
2) Cube de 7,5 cm de côté
300g
3) Cube de 5 cm de côté
100g
4) Cube de 2,5 cm de côté
5) Cylindre de 10 cm de diamètre et de 12 cm de haut
6) Cylindre de 4 cm de diamètre et 12 cm de haut
7) Tourillon de 8 mm de diamètre et 16 cm de haut
8) Balle de tennis
9) Briquet électronique
10) Bille de 25 mm de diamètre

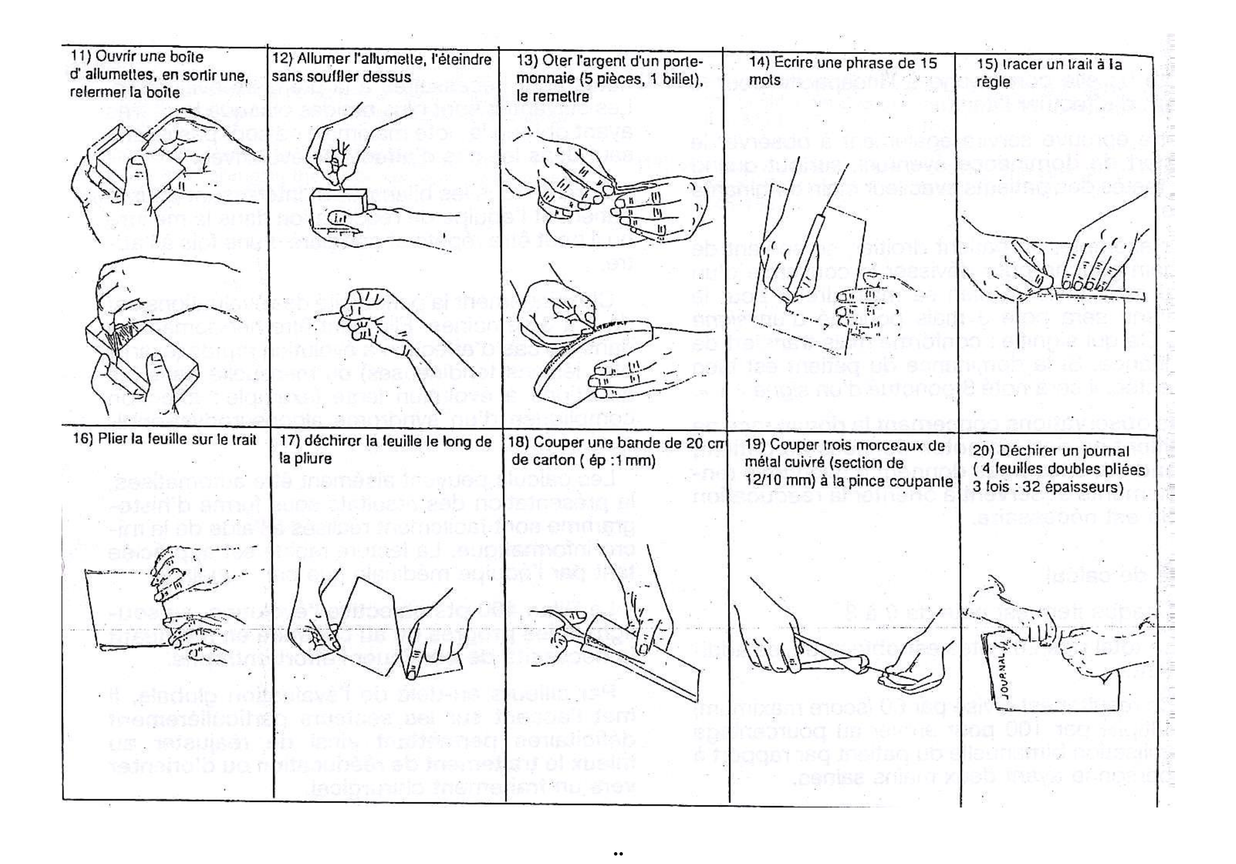


Trois objets nécessitent une action avant de les transporter sur le plan en hauteur : le briquet qu'il faut allumer, la clef qu'il faut engager dans la serrure, la déverrouiller, l'ouvrir et la retirer, et la cruche à moitié remplie d'eau (500 ml) dont il faut verser la moitié en pronation l'autre en supination dans un verre.
La cotation se fait de 0 à 3 :
Note 3 : la prise d'objet et le déplacement sont conformes ou convenables. Convenable quand le patient présente une séquelle définitive (amputation, arthrodèse..) et l'on admet les compensations.
Note 2 : la prise d'objet est conforme ou convenable mais le déplacement est lent ou imprécis ou la prise d'objet est disharmonieuse, imprécise et le déplacement conforme ou convenable.
Note 1 : la prise d'objet est conforme ou convenable mais le déplacement est incorrect (effectué avec compensations bien qu'il pourrait l'être correctement) ou la prise est incorrecte (effectuée avec exclusion ou compensation, bien qu'elle pourrait l'être correctement) mais le déplacement est conforme ou convenable.
Note 0 : l'action ou le transport sont impossibles.
En cas de dépassement du temps (> à une minute) le score n'est pas modifié mais cela est noté par le thérapeute.
Le calcul du score est la somme des notes (note maximale 60) divisée par 60 et multipliée par 100 pour obtenir le pourcentage d'utilisation mono manuelle de la main lésée par rapport à la main saine.

<u>EPREUVE 4 : FONCTION BI MANUELLE</u>.

On teste la fonction bi manuelle à l'aide de 20 tâches de la vie quotidienne. Cette épreuve n'est pas chronométrée et le patient ne reçoit pas de consignes.
Les 20 épreuves sont les suivantes (cf. schémas).


1) Couper 3 petits morceaux de pâte à modeler
2) Ouvrir une bouteille, la refermer : capsule plastique sertie
3) Ouvrir un bocal, le refermer : capsule vissée
4) Dévisser un écrou de 20 mm, le revisser complètement
5) Dévisser un écrou de 4 mm, le revisser complètement
6) Défaire 3 boutons
7) Rattacher les boutons
8) Dénouer 3 lacets
9) Les renouer
10) Enfiler une aiguille


11) Ouvrir une boîte d' allumettes, en sortir une, refermer la boîte
12) Allumer l'allumette, l'éteindre sans souffler dessus
13) Oter l'argent d'un porte-monnaie (5 pièces, 1 billet), le remettre
14) Ecrire une phrase de 15 mots
15) tracer un trait à la règle
16) Plier la feuille sur le trait
17) déchirer la feuille le long de la pliure
18) Couper une bande de 20 cm de carton (ép :1 mm)
19) Couper trois morceaux de métal cuivré (section de 12/10 mm) à la pince coupante
20) Déchirer un journal (4 feuilles doubles pliées 3 fois : 32 épaisseurs)
JOURNAL

La cotation se fait de 0 à 3 :
Note 3 : l'item est effectué de façon conforme ou convenable. Compensations admises comme l'épreuve 3.
Note 2 : l'item est effectué en entier mais avec lenteur et imprécision.
Note 1 : l'item est effectué en partie seulement ou l'item est effectué en entier mais avec exclusion ou compensation (bien que les capacités nécessaires au patient pour bien le réaliser soient présentes).
Note 0 : incapacité pour le patient d'effectuer l'item.
Cette épreuve sert à vérifier la dominance du sujet et son éventuel transfert sur la main d'appoint si la main lésée est dominante.
Pour le calcul c'est la même chose que l'épreuve 3. Le résultat est exprimé en fonction d'une personne ayant toute sa capacité bi manuelle.

La réalisation du bilan demande 30 minutes pour un thérapeute entraîné. Les évaluations ultérieures sont plus rapides car on ne refait pas les items ayant obtenus la note maximale. On peut le répéter toutes les 2, 3 ou 4 semaines en fonction du potentiel évolutif de la maladie en cause.
Le calcul peut être automatisé permettant un gain de temps et une lecture directe facile. Pour le calcul du score global du bilan, on fait la somme des résultats obtenus dans chacune des quatre épreuves et on divise par 4 pour avoir un résultat sur 100, qui est le pourcentage d'utilisation fonctionnelle de la main du patient.
L'intérêt est aussi de mettre en exergue les secteurs déficitaires permettant d'ajuster le traitement au mieux.

IV.2.7.2 Données statistiques sur le bilan 400 points.

Le bilan 400 points a été bien étudié sur le plan statistique dans deux études publiées en 1993 et 1996 (31,58). On ne peut que saluer une telle volonté de clarté de la part des auteurs du bilan 400 points.

Première étude (31) : elle a étudié la fiabilité inter observateur entre trois évaluateurs, deux ergothérapeutes, dont un connaît déjà le bilan 400 points et un kinésithérapeute.

Matériel et méthodes :

85 patients ayant une lésion du membre supérieur, 50 hommes et 35 femmes, entre Avril 1990 et Juillet 1990.
L'ergo 1 est l'ergothérapeute non habitué au bilan. L'ergo 2 est l'autre ergothérapeute. L'ergo 1 et le kinésithérapeute ont eu un apprentissage du bilan avec une projection audiovisuelle de vingt minutes et la lecture d'un guide succinct.
Le bilan est réalisé une fois devant les trois juges qui notent séparément, en même temps et sans aucune discussion entre eux.
La durée moyenne du bilan est de vingt minutes.

Méthodes statistiques :

La comparaison des cotations des deux juges est faite en analysant la distribution des différences et en calculant l'écart moyen. Pour les différentes activités de chaque épreuve,

l'accord entre deux juges est défini par le **coefficient kappa pondéré Kp.** Pour la comparaison des trois juges, il est utilisé un **coefficient kappa généralisé**, non pondéré, défini à partir du coefficient kappa simple.
Pour les scores des épreuves et de l'ensemble du bilan le **coefficient de corrélation intra classe et le coefficient linéaire** sont utilisés.
L'analyse de l'influence de l'âge, du sexe et du diagnostic du patient sur les différences de cotations est réalisée à l'aide de la **méthode de régression linéaire multiple pas à pas descendante et du coefficient de corrélation multiple.**

Résultats :

Le profil de la population montre une moyenne d'âge de 41 ans, une atteinte du membre dominant dans 47 % des cas. L'atteinte est orthopédique dans 60 % des cas et neurologique dans 22 % des cas.

Pour les épreuves 1 et 3 qui concernent les activités mono manuelles, l'accord entre juges est parfait pour le membre sain. L'analyse s'intéresse donc au côté lésé.
Dans l'ensemble, les cotations des deux juges non habitués au bilan (ergo 1 et kinésithérapeute) sont comparables.
L'ergo 2 habitué au bilan donne des cotations inférieures à celles des autres juges. Mais c'est entre l'ergo 2 et le kinésithérapeute que les différences sont les plus visibles.

Pour l'épreuve 1, la concordance entre juges est bonne à très bonne avec des coefficients kappa pondéré supérieurs à 0.70 à 0.80 pour de nombreuses activités.
Pour l'épreuve 3, la concordance est aussi très bonne, supérieure à 0.70.
Pour l'épreuve 4, la concordance est moyenne. Elle est très bonne pour 6 activités, bonne pour 12 activités et mauvaise pour 4 activités.

Entre les divers juges, la différence entre les scores est assez faible. Pour l'épreuve 1, les deux ergothérapeutes ont donné des cotations similaires, significativement inférieures à celles du kinésithérapeute. Pour les épreuves 3 et 4 les cotations de l'ergo 2 sont significativement inférieures à celles des autres juges, lesquels ont donné des cotations semblables.
La concordance entre ergo 1 et kinésithérapeute est très bonne pour toutes les épreuves. Les coefficients de corrélation intra classes et les coefficients de corrélation linéaire sont tous supérieurs à 0.90. Pour l'épreuve bi manuelle, la concordance entre les ergothérapeutes est bonne (0.68 et 0.82) et moyenne entre ergo 2 et kinésithérapeute (0.51 et 0.73).

L'analyse en régression linéaire multiple a montré que le diagnostic du patient et le fait que la lésion concerne ou non le membre dominant n'interviennent pas significativement sur la différence de cotations pour toutes les épreuves 1, 3 et 4, et pour le score global, quels que soient les juges considérés.

Discussion :

Les activités de l'épreuve 1 sont simples et bien définies ce qui explique que la cotation de cette épreuve ne pose en général pas de difficultés ce qui se retrouve dans les résultats entre les différents juges.
Les activités de l'épreuve 3 comportent l'étude de deux fonctions, prise de l'objet et déplacement, ce qui peut entraîner quelques différences dans la cotation. Les différences de cotation observées dans cette épreuve peuvent aussi être dues au fait que l'épreuve est

chronométrée. L'exécution des tâches est rapide ce qui peut entraîner pour un juge non habitué une cotation inappropriée. C'est dans cette épreuve que les juges peuvent mettre en évidence les mouvements anormaux, les exclusions et les compensations.
Les activités de l'épreuve 4 font intervenir de multiples fonctions des deux membres ce qui explique certainement une difficulté supplémentaire pour la cotation. Dans cette épreuve, la définition du geste standard est difficile et peut varier selon les juges. La différence entre note 1 ou 2 ou 3 est parfois difficile. En effet, le geste habituel du patient peut être différent du geste standard attendu sans pour autant être un mauvais geste ou un geste avec compensation. De même, la précision est parfois difficile à coter et il faut accepter une certaine tolérance si le geste est correct.
Après discussion avec les juges et un essai sur un patient, les explications avancées ont été confirmées.
Pour améliorer les cotations, il a été décidé en plus de l'apprentissage par la projection audio visuelle, de faire une formation rapide précisant :

- La définition des gestes standards et la position de référence des doigts.
- La manière d'aborder le patient en lui montrant le geste standard à suivre, au lieu de le laisser choisir un geste inadéquat.
- Le repérage dès le départ (et à la suite de l'épreuve 1) des anomalies décelables et l'identification des activités des épreuves 3 et 4 qu'elles peuvent influencer de manière à mieux cibler l'observation du juge.
- La tolérance sur la précision d'exécution et l'exclusion par exemple des doigts non nécessaires pour l'activité considérée.

En conclusion :
Le bilan 400 points a une bonne fiabilité inter observateurs globale. Celle-ci serait certainement bien meilleure si chaque évaluateur avait un entraînement préalable à la cotation de ce bilan. On peut regretter dans cette étude qu'il n'y ait pas eu une étude de la fiabilité inter observateurs entre deux juges entraînés. On peut aussi regretter l'absence d'étude de la fiabilité intra observateur et de la fiabilité test retest.

Deuxième étude (58) : elle s'est attaché à montrer l'unidimensionnalité du bilan 400 points. Dans un bilan comme le 400 points, il existe un grand nombre d'items de nature variée. Il faut donc identifier ou vérifier au préalable les dimensions cibles de l'outil, déterminer la position de chaque item dans ces dimensions et enfin examiner l'unidimensionnalité de l'outil c'est-à-dire vérifier qu'il ne représente qu'un seul concept ou dimension. Ce calcul se fait à l'aide de méthodes statistiques complexes.

Matériel et méthodes :

194 sujets ayant une lésion du membre supérieur vus sur une année entière à Nancy.
L'outil évalué est le bilan 400 points qui comprenait à l'époque 67 épreuves. Pour l'analyse statistique les épreuves sont cotées 1 activité normale ou 0 activité anormale.
Deux méthodes statistiques ont été utilisées pour analyser l'unidimensionnalité de l'instrument :

- L'analyse en composantes principales (ACP) :

Un instrument peut être considéré comme unidimensionnel si la variance expliquée par le premier facteur atteint un seuil entre 20 et 40 % de la variance totale et que le rapport de la

première sur la deuxième valeur propre est supérieur à 2. Le seuil minimal adopté pour retenir une corrélation avec une composante principale est de 0.30.
Le coefficient alpha de Cronbach est utilisé pour mesurer l'homogénéité de l'instrument donc la cohérence des items de l'instrument. Une fiabilité de 0.80 est nécessaire pour conclure à une bonne cohérence des items.

- La méthode basée sur le modèle de Rasch permet de calculer l'échelle de difficulté des tests fonctionnels et vérifier l'unidimensionnalité de l'outil. Cette méthode a déjà été expliquée au chapitre II.

Résultats :

Les patients étaient pour 75% des hommes, la moyenne d'âge était de 36 ans et le membre dominant était atteint dans 52 % des cas.

Atteinte du membre dominant :

- Analyse en ACP : les trois premiers facteurs principaux expliquent respectivement 44%, 10% et 4% de la variance totale. Le rapport entre la première valeur et la deuxième valeur était de 4,46. Les 67 items ont présenté des corrélations positives avec la première composante principale et seulement 6 items avaient des coefficients de corrélation inférieurs à 0.30. Les items de l'épreuve 3, 4 et dans une moindre mesure 1 étaient bien représentés sur ce premier facteur. Le deuxième facteur opposait l'épreuve 1 dont les items présentaient des coefficients de corrélation positifs à l'épreuve 3 dont les items présentaient des coefficients de corrélation négatifs. Quatre des six items de l'épreuve 2 étaient bien représentés sur le troisième facteur principal. Le coefficient alpha de Cronbach total était de 0.978.
  L'analyse en ACP a suggéré l'unidimensionnalité de l'instrument en cas d'atteinte du membre dominant.
- Modèle de Rasch : les différents items des quatre épreuves étaient relativement bien répartis sur cette échelle de difficulté. Cette analyse a confirmé l'hypothèse d'unidimensionnalité de l'instrument avec seulement 5 items déviants sur 67 (supination, pronation, adduction du pouce, écartement du pouce, pinces latérales contre résistance).

Atteinte du membre non dominant :

- Analyse en ACP : les trois premiers facteurs principaux expliquaient respectivement 35 %, 8% et 5 % de la variance totale. Le rapport entre la première valeur propre et la deuxième valeur propre était de 4,37. Neuf items ont présenté des coefficients de corrélation inférieurs à 0.30 avec la première composante principale (items de l'épreuve 2, pronation, supination et écrire une phrase de 25 mots). La contribution des items aux trois premiers facteurs principaux était semblable à celle obtenue dans le cas d'atteinte du membre dominant.
  Le coefficient alpha de Cronbach était similaire à celui observé dans l'atteinte du membre dominant.

- Modèle de Rasch : seul l'item 60 (écrire une phrase de 25 mots) différait de manière importante entre les deux types d'atteinte. Mais aucun item n'a été rejeté. L'hypothèse d'unidimensionnalité est aussi confirmée ici.

Analyse sur échantillon complet :

- Analyse en ACP : sur 173 sujets, elle a mis en évidence une variance expliquée par les trois premiers facteurs principaux de respectivement 41%, 8% et 4% de la variance

totale. Le rapport entre la première valeur propre et la deuxième valeur propre était de 4,68. Cinq items ont présenté des coefficients de corrélation inférieurs à 0.30 avec la première composante principale (items épreuve 2, pronation, supination).
Le coefficient alpha de Cronbach était de 0.976.

- Modèle de Rasch : seul l'item 16 (supination) a été rejeté. L'instrument est unidimensionnel.

Discussion :

Les résultats de l'ACP sont en faveur de l'unidimensionnalité de l'outil car:

- La variance expliquée par le premier facteur principal est élevée (> 35%).
- Peu d'items (9 au plus) sont à faible coefficient de corrélation avec ce facteur.
- La valeur du rapport entre les deux premières valeurs propres est élevée.

Mais, l'épreuve 2 est mal représentée sur le premier facteur et les items pronation et supination étaient mal représentés sur les trois facteurs.
La cohérence de l'outil est bonne avec un coefficient alpha de Cronbach toujours supérieur à 0.975. Cette valeur atteste de l'unidimensionnalité de l'instrument mais peut aussi traduire une redondance des items.
Le modèle de Rasch a montré que la seule différence entre membre dominant ou non était une difficulté supérieure pour écrire une phrase de 25 mots, ce qui est logique. Il a conclu aussi à l'unidimensionnalité de l'instrument. Il a permis aussi de hiérarchiser les épreuves. Dans l'ensemble, l'épreuve 3 est plus difficile que la 4. Seul le membre lésé est utilisé dans l'épreuve 3 alors que dans l'épreuve 4 le membre non lésé peut aider.

IV.2.7.3 Conclusions :

Nous avons longuement étudié le bilan 400 points. Il s'agit, il faut le signaler d'un des deux seuls bilans d'évaluation fonctionnelle de la main publiés en français. Malheureusement, il n'a pas été publié en anglais ce qui contribue certainement au fait qu'il ne soit pas connu et utilisé dans les pays anglo-saxons. Pourtant je pense qu'il s'agit d'un très bon outil d'évaluation de la main, très complet, à la fois descriptif et évaluatif. Il a été bien étudié, nous l'avons vu, sur le plan statistique avec une bonne fiabilité inter observateurs, une validité de contenu et une cohérence interne très bonne. Les principaux défauts de cet outil sont l'absence de données publiées sur la fiabilité intra observateur et test retest, sur la validité de construction et de critères, sur l'absence de données sur la sensibilité au changement et le fait qu'il faille pour l'évaluateur un apprentissage assez complet pour coter de manière juste les items. J'ajouterais aussi que l'outil tel qu'il a été publié en 1996 comporte quelques modifications par rapport aux items testés dans les études de 1993 et 1996, et que cette nouvelle version n'a pas fait l'objet de publications complémentaires depuis. Néanmoins, je pense que les qualités l'emportent sur les défauts et ce bilan 400 points devrait faire partie d'une évaluation globale de la main, telle qu'elle peut être pratiquée en centre spécialisé. Le tableau ci-dessous résume les caractéristiques du bilan.

PROPRIETES	BILAN 400 POINTS
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs et qualitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Observation
4) Interprétation des résultats	Analyse des tâches, pourcentage fonctionnel
5) Spécificité	Populations variées
6) Disponibilité	
Préfabriqué	Non
Domaine public	Oui
Langue de publication	Français
Coût	? (pas cher)
7) Temps demandé	
Administration	30 -45 minutes
Entraînement du testeur	Oui important
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	Français
Application dans une autre culture	Oui
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	Non
2) Equipement	Non préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	Oui la plupart du temps > 0.8
Intra observateur	?
Test-retest	?
Cohérence interne	Cronbach alpha > 0.975
3) Validité	
Contenu	Oui outil unidimensionnel
Construction	Pas étudiée
Relation avec AVQ	
Relation avec statut physique	
Distinction des groupes	
Critères	Pas étudiée
Sensibilité au changement	Pas étudiée
4) Normes	Non la norme c'est la main saine.

Tableau n°9: caractéristiques du bilan 400 points.

IV.2.8 Le test d'évaluation des membres supérieurs des personnes âgées (TEMPA).

Ce test a été publié la première fois en 1991 par une équipe canadienne de Sherbrooke dans la province de Québec. Puis en 1993 le manuel d'administration est édité. En 1993 et 1994 la fiabilité puis la validité ont été publiées. Enfin, en 1995 les normes du test ont fait l'objet d'un article (26).

IV.2.8.1 Description du test.

Le design de ce test a été conçu pour avoir des tâches très proches de la vie quotidienne des personnes âgées, uni ou bi latérales, et pour s'intéresser en plus du critère temps à la qualité de la tâche effectuée.
Ce test est standardisé et comprend 9 tâches représentatives de la vie quotidienne.
5 tâches sont bi manuelles :

- Ouvrir un pot de café et prendre une cuillère à café.
- Ouvrir une serrure.
- Ouvrir une boîte à pilules.
- Ecrire sur une enveloppe et coller un timbre.
- Mélanger et passer des cartes à jouer.

4 tâches sont uni manuelles :

- Prendre et transporter un pot.
- Prendre un pichet et remplir un verre.
- Manipuler des pièces de monnaie.
- Prendre et transporter de petits objets.

Deux mesures sont effectuées : la force de préhension au Jamar et l'endurance musculaire isométrique relative (Jamar + chronomètre).

Chaque tâche est précédée de consignes particulières et d'une démonstration. Le patient peut l'exécuter une fois avant le test chronométré.
Tous les tests sont chronométrés. La durée de passation du test est de 15 à 40 minutes.

Chaque tâche est mesurée selon trois critères :

- La vitesse d'exécution : chronomètre.
- L'évaluation fonctionnelle : avec une échelle à 4 niveaux.

0 : la tâche est faite avec succès sans hésitation ou difficulté.
1 : quelques difficultés avec la tâche.
2 : grande difficulté avec la tâche.
3 : le patient ne peut faire la tâche même avec de l'aide.

- L'analyse de la tâche quantifie les difficultés éprouvées par les patients selon 5 dimensions : force, mobilités, précision des mouvements grossiers, précision des mouvements fins et préhension.

IV.2.8.2 Données statistiques du TEMPA.

Ces données sont tirées de deux articles (26, 77).
Fiabilité inter observateurs : elle a été testée par deux testeurs en simultané sur 25 patients. Les coefficients de corrélation intra classe étaient supérieurs à 0.9 pour les scores d'évaluation fonctionnel et d'analyse de la tâche. Mais, seule une étude a été faite et le nombre de sujets testés est très faible.

Fiabilité intra observateur : pas de données.

Fiabilité test retest : dans une étude de 29 patients, avec un intervalle de retest de 3 à 7 jours, les coefficients de corrélation intra classe étaient supérieurs à 0.77 sauf pour la mesure d'endurance de la force. Mais cette fiabilité test retest n'a pas été étudiée chez les sujets sains et là aussi l'échantillon de sujets testés est faible.

Cohérence interne : pas de données.

Validité de construction : le TEMPA a été comparé à une mesure d'activités de la vie quotidienne (Functional autonomy measurement system) et des corrélations modérées à fortes ont été trouvées dans deux études entre 0.45 et 0.74. Les corrélations les plus fortes étaient avec les activités bi manuelles du TEMPA.
Pour étudier la relation entre le TEMPA et les mesures cliniques, le TEMPA a été comparé à l'ARA et au box and block test chez 104 patients. Il existe une forte relation dans la direction espérée (corrélations > 0.9) entre les résultats du TEMPA (analyse des tâches et évaluation fonctionnelle) et les scores de l'ARA et une corrélation modérée (> 0.7) entre TEMPA et box and block test.

Validité de critères : pas étudiée.

Sensibilité au changement : pas étudiée.

Normes : elles ont été établies chez 360 sujets âgés de plus de 60 ans dans une étude randomisée bien conduite (26). Elles sont valables pour chaque sexe et pour trois catégories d'âge et n'intéressent que le temps mis pour faire les différentes tâches. Le temps augmente avec l'âge de façon significative. L'âge, la perception de son état de santé, le niveau d'activité sont les meilleurs prédicteurs de la performance du test. La principale limitation de cette étude est que 97% des sujets étaient droitiers.

IV.2.8.3 Conclusions.

Le TEMPA est un bon instrument avec beaucoup de qualités. Il inclut des résultats qualitatifs à la différence de nombre d'autres outils, les tâches sont en relation avec les AVQ, sont bilatérales et il existe trois méthodes d'interprétation des résultats. Ces propriétés psychométriques gagneraient à être précisées et complétées mais si on le compare à nombres d'autres outils, ces propriétés sont meilleures globalement. Sa limitation majeure est qu'il est spécifique des sujets âgés bien que les tâches ne le soient pas. S'il fallait qu'il soit appliqué à d'autres populations les propriétés psychométriques devraient être à nouveau étudiées dans ces autres populations. Une de ses qualités est qu'il existe en français comme le bilan 400 points. Le tableau ci-dessous résume les caractéristiques de ce test.

PROPRIETES	TEMPA
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs et qualitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Observation
4) Interprétation des résultats	Tâches, évaluation fonctionnelle, normes.
5) Spécificité	Sujets âgés > 60 ans
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	Français et Anglais
Coût	340 $
7) Temps demandé	
Administration	15-40 minutes
Entraînement du testeur	Minimal
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	Français et anglais
Application dans une autre culture	Oui
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	Oui
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif, large champ
2) Fiabilité	
Inter observateurs	En partie > 0.9 dans une étude (25 patients)
Intra observateur	?
Test-retest	> 0.77 dans une étude sauf pour endurance
Cohérence interne	?
3) Validité	
Contenu	Oui
Construction	
Relation avec AVQ	Oui
Relation avec statut physique	Oui
Distinction des groupes	Pas étudiée
Critères	Pas étudiée
Sensibilité au changement	Pas étudiée
4) Normes	Oui pour la vitesse Que pour les + de 60 ans.

Tableau n°10: caractéristiques du TEMPA (d'après Rudman (77)).

IV.2.9 Le Sequential Occupational Dexterity Assessment (SODA).

Ce test a été mis au point par une équipe néerlandaise en 1995 et publié en anglais en 1996 (91) pour la version longue et en 1999 pour la version courte (90).
C'est un test de dextérité basé sur des tâches de la vie quotidienne et validé chez des patients ayant une PR.

IV.2.9.1 Le SODA version longue (91).

Matériel et méthodes :

109 patients ayant une PR, avec une moyenne d'âge de 54,5 ans et une durée moyenne de la maladie de 13 ans. Ces patients ont été testés une fois pour établir la cohérence interne et la validité du test.
Un autre groupe de 22 patients ayant une PR traitée par sels d'or a été utilisé pour établir la fiabilité test retest à 14 jours d'intervalle.
Un troisième groupe de 14 patients ayant une PR qui devaient être opérés de la main a été utilisé pour mesurer la sensibilité au changement (test avant chirurgie et 6 semaines après).
Un dernier groupe de 6 patients a été utilisé pour tester la fiabilité inter observateur (tests par trois évaluateurs différents).
Pour les associations entre variables, le coefficient de corrélation de Pearson a été utilisé.
D'autres informations ont été collectées : VS, nombre d'articulations douloureuses et gonflées, mesures des amplitudes articulaires, force au Jamar, score DAS (Disease Activity Score) d'activité de la PR, auto questionnaire d'évaluation de la dextérité et la douleur sur une EVA.

Résultats :

- Construction du SODA :

Le groupe de 109 patients a été testé dans un premier temps sur 23 tâches standardisées qui avaient été jugées par les patients comme importantes dans la vie de tous les jours. La cotation de ces tâches était : 0= incapable de faire la tâche, 1= capable de faire la tâche d'une manière différente, 4= capable de faire la tâche comme décrite.
Le patient devait coter la tâche effectuée : 0= très difficile, 1= quelques difficultés, 2= pas de difficultés. Une combinaison des deux scores représentait le score de chaque tâche.
De ces 23 tâches, 12 ont été sélectionnées car elles étaient sensibles à la variation et ces 12 tâches représentent le SODA.

1. écrire une phrase.
2. prendre une enveloppe.
3. ramasser des pièces.
4. prendre le combiné téléphonique et le mettre sur l'oreille.
5. dévisser le bouchon d'un tube de dentifrice.
6. mettre du dentifrice sur la brosse à dents.
7. manipuler une cuillère et un couteau.
8. boutonner une chemise.
9. dévisser une grande bouteille.
10. mettre de l'eau dans un verre.
11. se laver les mains.
12. se sécher les mains.

Toutes ces tâches sont décrites précisément dans un manuel d'utilisation du test. Pour 6 de ces 12 tâches, seule une main est notée alors même que les deux mains sont utilisées. Les 6 autres tâches sont notées pour chaque main séparément. Le score peut aller de 0 à 108. La durée d'administration est de 15 minutes en moyenne.

- Fiabilité du SODA :

La cohérence interne évaluée par le Cronbach alpha est de 0.91, ce qui est excellent.
La fiabilité test retest par le coefficient de corrélation de Pearson est de 0.93.
La sensibilité au changement testée chez 14 patients a montré un score SODA à 70,5 avant l'intervention et à 88,4 après l'intervention. Ceci prouve la sensibilité au changement du test apprécié par la méthode du t-test avec $t=3.65$ ($p< 0.01$).
La fiabilité inter observateurs sur seulement 6 patients a été appréciée par le coefficient de concordance W de Kendall $W= 0.78$ ($p<0.01$).

- Validité du SODA :

Les scores SODA ont été corrélés aux nombreuses variables recueillies. La première corrélation significative a été trouvée entre le score SODA et la durée de la maladie. Plus la maladie dure depuis longtemps plus le score SODA est élevé.
Les amplitudes articulaires du poignet, la mobilité des doigts et la force étaient toutes corrélées avec la dextérité donnée par le score SODA. La corrélation globale est de 0.71. Le carré de cette corrélation multiple donne le pourcentage de variation du score SODA qui est obtenue par la mesure d'autres variables. Une évaluation soigneuse de l'altération de chaque main explique seulement 51 % de la variation du score SODA.
La douleur générale (VAS) n'explique que 10 % de la variation dans le score SODA. Par contre, la douleur rapportée par le patient durant le test explique 28 % de la variation du score.
La corrélation entre le score SODA et l'auto questionnaire rempli par le patient est de 0.65. Cela veut dire que seulement 42 % de la variation est partagée entre la dextérité mesurée et celle auto évaluée par le patient.
Discussion :

Le SODA est un instrument fiable. La fiabilité inter observateurs doit être confirmée sur des échantillons plus importants.
Le SODA mesure l'incapacité de la main et non pas l'altération de la main. L'altération de la main mesurée par la force et les mobilités n'explique que 50 % de la variation du score. Cela veut dire que des patients ayant le même degré d'altération de la main vont avoir une incapacité qui varie considérablement.
En ce qui concerne la différence entre la dextérité mesurée et celle rapportée par les patients, cela veut dire que certains patients avec le même score auto évalué peuvent avoir des scores très différents au SODA. Cette différence pourrait s'expliquer par l'utilisation par les patients d'aides techniques à domicile.

IV.2.9.2 Le SODA court ou short SODA (SODA-S).

En 1999, les mêmes auteurs ont publié cette version courte du SODA (90). Ils ont repris le même collectif de 109 patients ayant une PR déjà testés dans la précédente étude. Ils ont été testés à nouveau à un an. Il ne restait alors plus que 94 patients. Sur les 12 tâches, 6 ont montrées des scores significativement signifiants entre t 0 et t 1 an. Ces 6 tâches constituent le SODA-S. Ce sont les suivantes :

1. prendre une enveloppe.
2. prendre des pièces de monnaie.

3. prendre le combiné téléphonique.
4. ouvrir le tube de dentifrice.
5. mettre le dentifrice sur la brosse à dents.
6. manipuler une cuillère et un couteau.

Il y a donc 3 tâches uni manuelles et 3 tâches bi manuelles.
Sur le plan statistique la consistance interne par le Cronbach alpha est de 0.85. La corrélation avec le SODA est de 0.93 à t 0 et de 0.94 à t 1 an. Les corrélations sont identiques à celles obtenues avec le SODA. La sensibilité au changement est bonne mais cela n'est pas surprenant car les items sélectionnés sont les plus signifiants.

L'avantage du SODA-S est de ne prendre que 10 minutes à faire tout en donnant les mêmes renseignements que le SODA.
Le SODA et le SODA-S sont capables de mesurer des différences entre les patients (fonction discriminative), grâce aux données en percentiles des scores. Ces tests ayant une bonne sensibilité au changement, ils peuvent être utilisés pour mesurer l'effet d'une thérapie. On peut comparer les scores pré et post traitement avec les déviations standards de ces deux tests. Un changement de 1 DS entre les deux mesures est un changement très substantiel, de 0,5 DS est relativement bon et 0,25 DS est un petit changement. Mais il est difficile de dire si ce changement est cliniquement pertinent.

IV.2.9.3 Conclusions.

Le SODA et le SODA-S sont des tests simples à utiliser, très bien étudiés sur le plan statistique et dont la conception globale est bonne. La fiabilité inter observateurs aurait mérité d'être mieux étudiée sur un échantillon plus important. La fiabilité intra observateur n'a pas été mesurée. Le principal défaut de cet outil est d'avoir été validé dans une population de PR et d'être valable actuellement uniquement dans cette population. Si l'on voulait utiliser cet outil dans une autre population plus variée de nouvelles études de fiabilité et validité seraient à faire. Dans l'état actuel on ne peut que recommander son utilisation mais uniquement dans une population de PR. De plus l'existence du manuel uniquement en hollandais limite son utilisation.
Le tableau ci-dessous résume les caractéristiques des deux tests.

PROPRIETES	SODA	SODA-S
UTILITE CLINIQUE		
1) Type des résultats	Qualitatifs	Idem
2) Type des tâches		
Représentatif des AVQ	Oui	Oui
Unilatéral	Oui	Oui
Bilatéral	Oui	Oui
3) Méthode d'administration	Observation	Idem
4) Interprétation des résultats	Tâches, évaluation fonctionnelle.	Idem
5) Spécificité	PR	Idem
6) Disponibilité		
Préfabriqué	Non	Non
Domaine public	Oui	Non
Langue de publication	Hollandais	Idem
Coût	200 $ pour le manuel d'utilisation	?
7) Temps demandé		
Administration	20 minutes	10 min
Entraînement du testeur	Minimal	Idem
8) Acceptabilité		
Description du but	Oui	Oui
Approprié à des adultes	Oui	Oui
Langue	Hollandais	Idem
Application dans une autre culture	Oui	Oui
STANDARDISATION		
1) Instructions		
Administration	Oui	Oui
Calcul du score	Oui	Oui
Interprétation	Oui	Oui
2) Equipement	Non préfabriqué	Idem
BUTS	Descriptif et évaluatif	idem
PROPRIETES PSYCHOMETRIQUES		
1) Construction des items	Descriptif.	Idem
2) Fiabilité		
Inter observateurs	Oui en partie mais échantillon faible	Idem
Intra observateur	Non testée	idem
Test-retest	Oui 0.93	0.85
Cohérence interne	Oui Cronbach alpha 0.91	
3) Validité		Idem
Contenu	Pas étudiée	
Construction		
Relation avec AVQ	Oui	Idem
Relation avec statut physique	Oui	Idem
Distinction des groupes	Oui	Idem
Critères	Pas évaluée	Idem
Sensibilité au changement	Oui	Oui
4) Normes	Non	Non
PERSPECTIVE DU PATIENT	Oui	Oui

Tableau n°11: caractéristiques du SODA et du SODA –S.

IV.10 Le Sollerman hand function test.

Il a été publié en 1995 par une équipe suédoise (82). Il semble être le fruit d'une longue réflexion débutée en 1977 (42). La première publication de Sollerman sous forme d'une thèse date de 1980. Il s'agit d'un test inspiré de l'UEFT de Caroll, mais adapté et basé sur les 7 plus fréquentes préhensions.

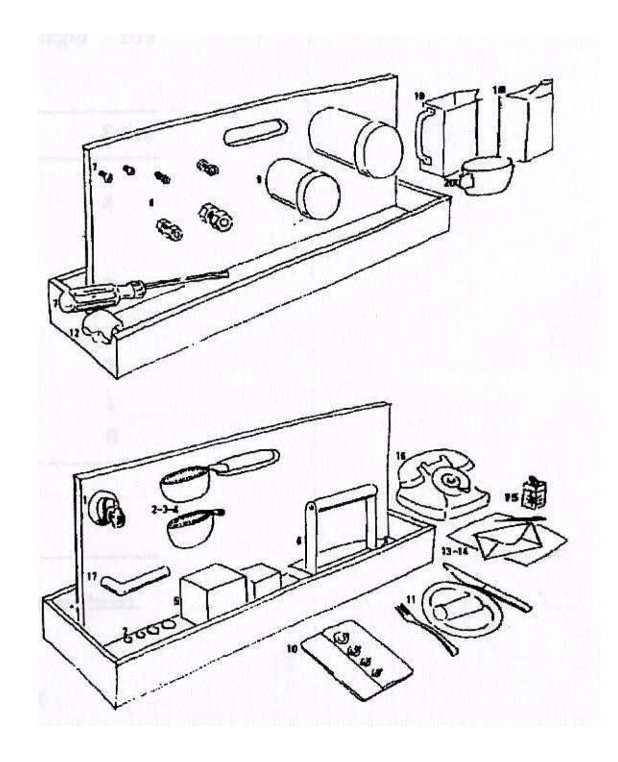
- **Description du test :**

  Ce test comprend 20 sous tests qui sont des activités quotidiennes.

  1. Prendre une clé, la mettre dans une serrure et tourner 90° (pince pulpaire et latérale).
  2. Prendre 4 pièces de monnaie et les mettre dans un porte monnaie sur un support (pince pulpaire).
  3. Ouvrir et fermer une fermeture éclair (pince pulpaire et latérale).
  4. Prendre les pièces dans le porte monnaie (pince pulpaire).
  5. Prendre deux cubes de bois et les mettre en hauteur sur une tablette (pince 5 doigts).
  6. Porter un fer à repasser de 3 Kg et le placer sur la tablette (prise à pleine main transverse doigts fléchis).
  7. Tourner deux vis avec un tournevis, un tour en supination (prise à pleine mains diagonale).
  8. Prendre 4 écrous de différentes tailles et les mettre sur les boulons appropriés (pince pulpaire, latérale et tri digitale).
  9. Dévisser deux couvercles de pots (prise sphérique).
  10. Déboutonner 4 boutons (pince pulpaire et latérale).
  11. Couper de la pâte à modeler avec un couteau et une fourchette (pince tri digitale et prise diagonale).
  12. Mettre un bandage élastique tubulaire sur l'autre main (pince latérale et à 5 doigts).
  13. Ecrire son nom (pince tri digitale).
  14. Plier une feuille de papier et la mettre dans l'enveloppe (pince latérale et prise 5 doigts).
  15. Mettre un trombone sur une enveloppe (pince pulpaire et latérale).
  16. Prendre le combiné téléphonique et le mettre sur l'oreille (prise à pleine main diagonale).
  17. Tourner un bouton de porte de 30° en supination (prise à pleine main transverse).
  18. Verser un litre d'eau d'un récipient dans un broc d'eau (prise à 5 doigts).
  19. Verser l'eau du broc dans une tasse (prise à pleine main transverse).
  20. Verser le contenu de la tasse dans le broc (prise pulpaire et latérale).

  Chaque épreuve est notée de 4 à 0 :
  4= Tâche bien effectuée avec la bonne préhension en moins de 20 secondes.
  3= Tâche effectuée avec un peu de difficulté, ou tâche faite entre 20 et 40 secondes, ou tâche faite avec une préhension un peu différente.
  2= Tâche effectuée avec grande difficulté ou tâche faite entre 40 et 60 secondes ou tâche faite mais avec une préhension incorrecte.
  1= Tâche partiellement faite en moins de 60 secondes.
  0= Tâche impossible à faire.

Les sujets normaux ont un score de 80 points pour la main dominante et de 77 à 79 pour la main non dominante.
Le matériel pour le test est monté sur un support. Le sujet est assis. Le test est administré une main après l'autre sauf les épreuves 11,14 et 15 qui nécessitent les deux mains.



- **Evaluation de la validité du test :** 40 patients représentant 47 mains ont été évalués. Les diagnostics étaient variés : PR, amputations digitales, blessures neurologiques, raideurs articulaires. Les patients devaient évaluer leur fonction de la main sur une EVA entre « pas de fonction de la main » et « fonction normale ». La corrélation entre le test de Sollerman et l'auto évaluation du patient était bonne (r=0.68). La façon de calculer cette corrélation n'est pas explicitée dans l'article.
- **Fiabilité :** elle a été calculée sur 10 patients. La fiabilité test retest et inter observateurs est de 0.98. Il n'y a pas plus d'explications dans l'article.
- **Etude sur des tétraplégiques :** 59 patients ont eu le test de Sollerman en pré opératoire. Ils devaient être opérés de leurs mains pour des transferts tendineux. Ils étaient classés en fonction de la classification de Giens qui est une classification des muscles moteurs restants. Dans le groupe de patients ayant une sensibilité tactile conservée, la corrélation entre le score de Sollerman et la classification de Giens était élevée (r=0.76 avec p<0.001).
- **Sensibilité au changement :** aucune donnée n'est disponible dans cette étude.

Si l'on se base sur cette étude, on ne peut pas dire que les qualités psychométriques du test de Sollerman soient bonnes et aient fait l'objet d'une étude extensive. Les résultats publiés ne sont pas très précis et ne concernent qu'un petit collectif. Le test en lui-même semble bien conçu et précis dans ses modalités d'administration, bien qu'il n'apporte pas de révolution par rapports aux tests antérieurs. Par contre, sa cotation est bien faite et plus facile à utiliser que

celle du bilan 400 points. Comme souvent, il a été repris dans des publications ultérieures dont quelques unes semblent intéressantes.
O'Connor (70), en 1999, a comparé le SODA et le test de Sollerman chez 25 patients ayant une PR. Ces patients ont aussi rempli le questionnaire HAQ. Il existe une association forte entre les deux tests. Les corrélations entre les deux instruments et le HAQ sont significatives. Ceci est plutôt en faveur d'une validité de construction du Sollerman correcte, car le SODA a de bonnes propriétés psychométriques (cf chapitre IV.2.9).
Limaye (57), en 2001, a évalué la main chez 35 patients hémodialysés depuis en moyenne 6 ans. L'évaluation comportait le test de Sollerman, le HAQ et diverses variables cliniques. Le score moyen du test de Solerman était 77, et 54 % des patients avaient un score inférieur à la normale qui est de 78 à 80. La corrélation était significative avec l'HAQ (r=0.66), la durée de l'hémodialyse, et la force au Jamar. Les scores de Sollerman étaient hautement corrélés entre main dominante et non dominante (r=0.69 avec $p<0.00005$). Le test de Sollerman semble dans cette étude bien refléter la fonction de la main des patients hémodialysés et pourrait être utilisé pour détecter l'altération de cette fonction, fréquente avec l'amylose des dialysés.
Sinha (80), en 2002, dans une étude prospective sur la chirurgie du Dupuytren chez 42 patients, a utilisé le Sollerman comme outil d'outcome. Le score moyen de Sollerman est passé de 71 en pré opératoire à 77 en post opératoire. Il existe dans cette étude une corrélation significative entre le degré de déformation et le score de Sollerman et aussi entre l'amélioration de la déformation après chirurgie et le score de Sollerman. Ceci est un argument pour la sensibilité au changement du test de Sollerman.
Draviaraj (27), en 2004, rapporte aussi une étude sur la chirurgie du Dupuytren chez 30 patients. Le test a été administré avant la chirurgie puis 6 et 12 mois après. Le score moyen du test était de 72,8 avant l'intervention et de 77,9 12 mois après la chirurgie. Il existait dans cette étude une corrélation significative entre la correction totale, la correction de la flexion de l'IPP et la fonction de la main à 6 et 12 mois. La corrélation avec la correction de la flexion de la MP n'était pas bonne.

En conclusion, ce test a probablement des qualités certaines, mais pour qu'il puisse être utilisé sans restriction, il faudrait que ses qualités psychométriques soient étoffées et qu'une autre étude s'attache à calculer à nouveau la fiabilité test retest, la fiabilité inter observateurs et intra observateur, la cohérence interne et la validité de construction. Le tableau ci-dessous résume ses caractéristiques.

PROPRIETES	SOLLERMAN TEST
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs et qualitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Observation
4) Interprétation des résultats	Tâches, évaluation fonctionnelle, normes.
5) Spécificité	non
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	Suédois et Anglais
Coût	?
7) Temps demandé	
Administration	20-25 minutes
Entraînement du testeur	Minimal
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	Suédois et Anglais
Application dans une autre culture	Oui
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	non
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif, large champ
2) Fiabilité	
Inter observateurs	0.98 dans une étude
Intra observateur	?
Test-retest	0.98 dans une étude
Cohérence interne	?
3) Validité	
Contenu	Oui a priori
Construction	
Relation avec AVQ	Oui a priori mais pas prouvée
Relation avec statut physique	?
Distinction des groupes	Pas étudiée
Critères	Pas étudiée
Sensibilité au changement	Pas étudiée de façon claire
4) Normes	Non pas en fonction de l'âge et du sexe. Norme globale

Tableau n°12: caractéristiques du Sollerman test.

IV.3 CONCLUSIONS.

Nous avons vu dans ce chapitre 10 tests de dextérité basés sur des activités de la vie quotidienne mis au point entre 1965 et 1996. L'étude de ces tests a permis de montrer que la plupart d'entre eux, même si le concept de base qui a conduit à leur construction était bon, souffrent d'insuffisance sur le plan statistique dans leur mise au point. Il faut reconnaître que depuis les années 1980 les outils sont de meilleure qualité.
De tous les outils mis au point, le bilan 400 points semble le plus abouti, même s'il existe des insuffisances que j'ai souligné plus haut. Il englobe outre la dextérité, la force et la mobilité globale des doigts et de la main. Il a l'énorme avantage pour nous francophones d'être en français et d'être bien standardisé. Il n'a malheureusement pas été publié en anglais ce qui aurait pu assurer son succès et des études plus nombreuses. Bien sur, il demande un peu de temps et semble difficile à faire dans une pratique privée d'ergothérapeute ou de rééducateur de la main. En centre, en combinaison avec les tests classiques de dextérité il trouve toute sa place. La difficulté, comme avec la plupart des tests, se trouve dans l'interprétation des scores. Il est en effet difficile de savoir à partir de combien de pourcentage d'amélioration il existe une significativité et surtout si cette différence est cliniquement pertinente pour le patient. Seules des études à grande échelle avec des pathologies homogènes et en association avec d'autres outils de mesure (questionnaire) permettraient de répondre à cette question. Ce bilan 400 points me paraît donc un outil indispensable à une évaluation fonctionnelle complète de la main. Il faut à mon avis le réserver plutôt à des cas complexes ou d'évolution longue, de préférence d'origine orthopédique ou neurologique périphérique.
Les deux autres tests que l'on peut retenir sont le TEMPA et le SODA. Ces deux outils ont de bonnes qualités psychométriques globales. Leur principal défaut est d'être des outils spécifiques d'une population donnée (personnes âgées et PR) et donc peu utilisables en l'état actuel dans d'autre populations.
Tous les autres tests n'ont en l'état actuel des connaissances pas suffisamment de qualités pour pouvoir être utilisés sans retenue.
J'ai fait un tableau récapitulatif de tous les tests avec les principales propriétés (**Tableau n°13** cf ci-dessous). Pour mémoire (chapitre II), un instrument descriptif doit avoir 4 critères psychométriques fondamentaux, un instrument évaluatif 6 et un instrument à la fois descriptif et évaluatif 8.
Tous ces tests devraient actuellement être utilisés en association avec d'autres outils et en particulier les questionnaires d'auto évaluation qui étudient d'autres dimensions du retentissement fonctionnel, non évaluées par ces différents tests. Nous allons au chapitre suivant voir quels sont les questionnaires utilisables.

TESTS	**UEFT**	**Functional Hand testing**	**JEBSEN**	**SHFE**	**ARA**	**400 Points**	**TEMPA**	**SODA/ SODA-S**	**SOLLER-MAN**
TEMPS	20 – 30 min.	?	15 min.	20 min.	5 à 10 min.	30 à 45 min.	15 à 50 min.	20 min./10 min.	20 à 25 min.
RESULTATS	Quantitatifs	Quantitatifs + Qualitatifs	Quantitatifs	Quantitatifs	Quantitatifs	Quantitatifs + Qualitatifs	Quantitatifs + Qualitatifs	Quantitatifs	Quantitatifs + Qualitatifs
BUTS	Descriptif + Evaluatif	Descriptif + Evaluatif	Descriptif	Descriptif + Evaluatif	Descriptif + Evaluatif	Descriptif + Evaluatif	Descriptif + Evaluatif	Descriptif + Evaluatif	Descriptif + Evaluatif
PROPRIETES PSYCHO-METRIQUES	0 à 1/8	0/8	1 à 2/4	1/8	3/8	4/8	5/8	5/8	3/8
NOMBRE D'ITEMS EVALUES	33	12	7	13	19 maximum	57	9	12 ou 6	20

V. LES QUESTIONNAIRES D'AUTO-EVALUATION.

V.1. Introduction.

C'est depuis les années 1990 que l'on s'intéresse beaucoup plus au point de vue des patients par le biais d'auto questionnaires d'évaluation. Nous avons vu au chapitre II que la perspective du patient doit faire partie de tout bon outil d'évaluation. Les cliniciens sont plus habitués à des mesures dites objectives (force, amplitudes articulaires, sensibilité, radiographies..) qu'à l'utilisation de ces auto-questionnaires. Ors, on le sait, ces mesures objectives sont souvent peu fiables et peu reproductibles. Il a même été montré que la fiabilité et la sensibilité au changement de certains questionnaires étaient supérieures aux données dites objectives (4). Bien sûr, le patient peut mentir et remplir le questionnaire de façon malhonnête, mais il peut aussi lors de l'examen clinique présenter des autolimitations ou bien donner des réponses fausses par exemple à l'examen de la sensibilité. Comme l'examen physique, le remplissage de ces questionnaires nécessite la coopération du patient. Pour la main et le poignet, la plupart de ces questionnaires ont été mis au point entre 1993 et 2004 et certains sont déjà devenus des outils très utilisés dans le monde de la chirurgie de la main et du membre supérieur en général. Leur mise au point répond à des critères stricts que nous verrons. J'ai sélectionné 9 outils :

- Trois pour le membre supérieur dans son ensemble, y compris la main et le poignet (DASH=Disabilities of the Arm, Shoulder and Hand, UEFS= Upper Extremity Function Scale, POS hand /arm= Patient Outcomes Of Surgery hand/arm).
- Un pour la main et le poignet (Michigan hand questionnaire=MHQ).
- Deux pour le poignet uniquement (PRWE= Patient Rated Wrist Evaluation et wrist outcome instrument).
- Un pour la main uniquement (ABILHAND).
- Deux questionnaires spécifiques du canal carpien (Boston questionnaire=BQ et le AMHFQ= Alderson-McGall Hand FunctionQuestionnaire).

Un dernier outil le HFS (Hand Function Scoring system) n'a pas été retenu car non validée (94).

Nous allons voir en détails ces outils toujours avec la méthode utilisée au chapitre IV pour vérifier si les qualités attendues sont présentes.

Nous ferons une synthèse des outils utilisables en pratique courante en association avec les tests classiques et l'examen clinique.

V.2. Le questionnaire DASH: Disabilities of the Arm, Shoulder and Hand.

V.2.1 Introduction.

La publication initiale du DASH a été faite en 1996 (41), par des auteurs canadiens et américains au nom d'un groupe de travail sur le membre supérieur appelé « Upper Extremity Collaborative Group= UECG ». Depuis sa publication, il y a presque 10 ans, ce questionnaire est reconnu, traduit et validé dans 14 langues différentes. Une quarantaine de publications en rapport avec ce questionnaire sont recensées. Il est devenu un outil de référence en matière de questionnaire générique sur le membre supérieur dans son ensemble. Nous allons passer en revue la mise au point du DASH, son mode d'emploi, ses caractéristiques psychométriques et son interprétation.

V.2.2 Développement du DASH.

Le DASH a répondu à des critères stricts et à des étapes précises pour sa mise au point. Il est apparu aux auteurs qu'il fallait pour les pathologies du membre supérieur mettre au point un outil évaluatif et descriptif de mesure, outil qui faisait défaut jusque là. Le processus est bien décrit dans la publication princeps et dans le manuel d'utilisation du DASH (41,83).
L'UECG est un groupe pluridisciplinaire qui comprend des kinésithérapeutes, des chirurgiens de la main, des rhumatologues, des rééducateurs, des ergothérapeutes et des épidémiologistes. L'UECG a travaillé en collaboration avec l'association des orthopédistes américains (AAOS) et l'institut du travail et de la santé de Toronto.
Le but de ce travail coopératif était de développer une mesure régionale d'outcome qui conceptualiserait le membre supérieur comme unité fonctionnelle unique. Davis et collaborateurs (24) ont montré en étudiant des items d'outils déjà existants, dont certains étaient des outils spécifiques d'une articulation, que ces outils incluent une proportion significative d'items (environ 80%) non relatifs à une articulation spécifique mais relatifs au membre supérieur dans son ensemble. Ceci est bien corrélé à la pratique clinique quotidienne, même s'il existe quelques biais dans cette étude.
Cet outil pourrait être utilisé pour l'évaluation de n'importe quelle articulation ou pathologie du membre supérieur, et pourrait être utile pour des comparaisons entre différentes pathologies du membre supérieur (descriptif) et pour juger de l'efficacité d'une intervention thérapeutique (évaluatif). Cette mesure s'adresserait à tous, sédentaires, sportifs, travailleurs, quel que soit l'articulation atteinte, quelle que soit la pathologie et que celle-ci soit aigue ou chronique. Le but du DASH est d'évaluer les symptômes et le statut fonctionnel en se centrant sur la fonction physique dans des populations variées de pathologies du membre supérieur. Dans le tableau ci-dessous les différentes dimensions du DASH sont résumées.

CONCEPT	DIMENSIONS	COMPOSANTS
SYMPTOMES		Douleur
		Faiblesse
		Raideur
		Picotements/Engourdissements
STATUT FONCTIONNEL	**PHYSIQUE**	Activités de la vie quotidienne
		Occupations ménagères/jardinage
		Courses (commissions)
		Activités récréatives
		Hygiène personnelle
		Habillage
		Manger
		Activité sexuelle
		Sommeil
		Sports/activités artistiques (Optionnels)
	SOCIALE	Famille
		Travail
		Socialisation avec amis
	PSYCHOLOGIQUE	Image de soi

Tableau n°14 : Concepts, dimensions et composants du DASH, d'après Hudak (41).

Trois étapes sont nécessaires et indispensables pour mettre au point un questionnaire et en particulier le DASH.

Etape 1 : sélection des items.
Cette étape a consisté à sélectionner par une revue de la littérature 13 échelles existantes sur le membre supérieur, générant une liste de 821 items. Les items sont ensuite classés selon la classification OMS : symptômes musculo-squelettiques, altération (impairment), incapacité (disability) et handicap.

Etape 2a : réduction initiale des items basée sur le jugement d'experts.
Elle se fait par opinion de trois experts de l'UECG qui ont réduit la liste des items à 177. Ces 177 items sont ensuite soumis à d'autres experts et à l'UECG pour étudier la validité de contenu des items. Le nombre d'items est alors réduit à 67.
Ensuite un questionnaire avec 67 questions a été formaté par le groupe et administré à titre de pré test à 20 patients. Le questionnaire comportait une échelle de réponses de 5 points (1 à 5) et la période choisie pour que le patient puisse se rappeler, était les 7 derniers jours. Les questions sont axées sur les capacités de l'individu à faire l'activité et le sujet doit s'auto évaluer sur une échelle à 5 degrés : aucune difficulté, difficulté légère, difficulté moyenne, difficulté importante et impossible.
Les patients étaient interrogés sur le contenu des questions, sur la compréhension des questions et sur l'échelle de réponse. 3 items supplémentaires sur l'image de soi ont été introduits portant le nombre d'items à 70. Les experts ont aussi estimé que le sport et les arts n'étaient pas bien représentés. Un module optionnel de 4 items a été rajouté.

Etape 2 b : réduction supplémentaire des items par des méthodes statistiques.

Un des buts du DASH était d'avoir un questionnaire assez court, facile à remplir avec approximativement 30 items. Pour cela deux stratégies statistiques peuvent être utilisées la stratégie psychométrique et la stratégie clinimétrique.
La stratégie clinimétrique est surtout utilisée en médecine clinique et se base sur les jugements des patients et des médecins. Elle a comme but de mesurer des phénomènes cliniques qui sont généralement connus pour contenir des caractéristiques ou des attributs du patient qui sont sans rapport avec ce que l'on veut mesurer. Les échelles qui mesurent de nombreux attributs sont appelées échelles hétérogènes.
La stratégie psychométrique, utilisée en psychologie et éducation, fait appel aux techniques mathématiques, et son but est de faire des échelles soit simples, soit multiples mais qui mesurent une seule caractéristique ou attribut du patient. Les échelles qui mesurent un seul attribut sont appelées homogènes.
Pour le DASH et pour être sur de choisir la bonne, les auteurs ont utilisé les deux approches.
L'étude a été faite sur 407 patients dans 20 centres différents au Canada, USA et Australie.
L'analyse statistique est très complexe et je ne la décrirais pas en détails.
Pour l'analyse psychométrique, le Cronbach alpha est de 0.975 pour les 30 items sélectionnés, ce qui est excellent et donne une très bonne cohérence interne. Une analyse en ACP (cf chapitre II) a été pratiquée sur les 70 items et un seul facteur explique 45% de la variance des 70 items ce qui confirme l'unidimensionnalité de cet outil et le fait que les 30 items sont bien représentatifs.
Pour l'analyse clinimétrique, qui a intéressée 76 patients, le questionnaire utilisé était le même avec en dessous de chaque question deux sous questions : comment cette activité était-elle pénible ? Et quelle est l'importance de l'activité pour le patient ? Pour l'analyse, les 30

items ayant le score d'importance le plus élevé ont été considérés comme les 30 items les plus importants.

Etape 2c : élaboration de la version finale du DASH à partir des deux analyses.

Les résultats ont été comparés par des experts de l'UECG et 30 items ont donc été sélectionnés plus 4 items pour le module optionnel sports/arts.
Une analyse statistique a donc été faite sur ces 30 items. Le coefficient de corrélation de Spearman a été calculé et une analyse en ACP a été pratiquée. Un premier facteur très fort, expliquant entre 50 et 60 % de la variance, a été mis en évidence. Les coefficients de corrélation dans ce modèle, pour chaque item sont entre 0.40 et 0.83 avec une grande majorité de coefficients supérieurs à 0.70.
La version finale du DASH a donc été acceptée en 1996. Elle comprend 30 items obligatoires : 21 items sur la fonction physique, 6 items sur les symptômes et 3 items sur la fonction sociale. Deux modules optionnels de 4 items chacun sur le sport/musique et sur le travail complètent le DASH.
Le module sur le travail est fait selon le même modèle que le module sport et a été rajouté pour différencier les patients qui pouvaient ne pas être gênés dans leurs AVQ, mais gêné dans une activité spécifique du travail.

La troisième étape de la validation du DASH concerne l'étude des caractéristiques psychométriques et fait l'objet du chapitre suivant.

V.2.3 Propriétés psychométriques du DASH.

Nous avons vu au chapitre II l'importance des propriétés psychométriques d'un outil d'évaluation. Le DASH étant à la fois un outil évaluatif et descriptif il doit avoir les qualités suivantes :

- Une fiabilité test retest élevée.
- Une cohérence interne forte.
- Une validité de contenu et de construction prouvée.
- Une bonne sensibilité au changement.

Toutes ces qualités ont été démontrées dans différentes études. Je ne me suis intéressé qu'aux études ayant rapport avec la main et/ou le poignet ou aux études plus globales sur le membre supérieur qui comportaient des patients ayant des pathologies de la main ou du poignet.

V.2.3.1 Fiabilité (10, 11,83).

La fiabilité test retest et la cohérence interne sont les deux caractéristiques importantes pour le DASH puisque les fiabilités inter et intra observateurs ne sont pas à mesurer en raison du caractère auto administré du questionnaire.
Cohérence interne : le Cronbach alpha a été calculé à 0.96 lors du développement du questionnaire (cf.supra). Cet excellent chiffre a été retrouvé lors du calcul du Cronbach alpha dans les versions étrangères validées du DASH. (cf. tableau ci-dessous).
Fiabilité test retest : elle a été étudiée par les auteurs du DASH dans une grande étude (11) portant sur 200 patients qui présentaient des pathologies variées du membre supérieur. Ces patients étaient évalués avant et 12 semaines après traitement. 56 patients ont refait le DASH entre 3 et 5 jours après le premier remplissage. Le coefficient de corrélation de Pearson était de 0.96 ce qui est excellent. Dans une étude sur des pathologies de coude, cette fiabilité a été

mesurée à 0.92. Cette excellente fiabilité test retest est aussi confirmée dans les versions traduites et validées.

VERSIONS DU DASH	CRONBACH ALPHA	TEST RETEST
Américaine (10,11)	0.96	0.96
Suédoise (7)	0.96	0.92 (67 patients délai 7 jours)
Espagnole (76)	0.95	0.97 (42 patients délai 7 jours)
Néerlandaise (92)	0.95	0.98 (50 patients délai ?)
Italienne (73)	0.90	0.89 (30 patients délai 5 jours)
Allemande version n°1(71,72)	0.96	0.90 (18 patients délai 10 jours)
Allemande version n°2 (35,36)	Pas disponible	Pas disponible
Canadien français (29)	0.94	Pas disponible
Chinoise (55)	0.94	0.77 (88 patients délai 7 à14 jours)
Française (28)	Pas disponible	Pas disponible

Tableau n°15 : fiabilité des différentes versions du DASH.

V.2.3.2 Validités.

La question essentielle est donc de savoir si le DASH mesure ce qu'il est censé mesurer ? (10, 83,84)

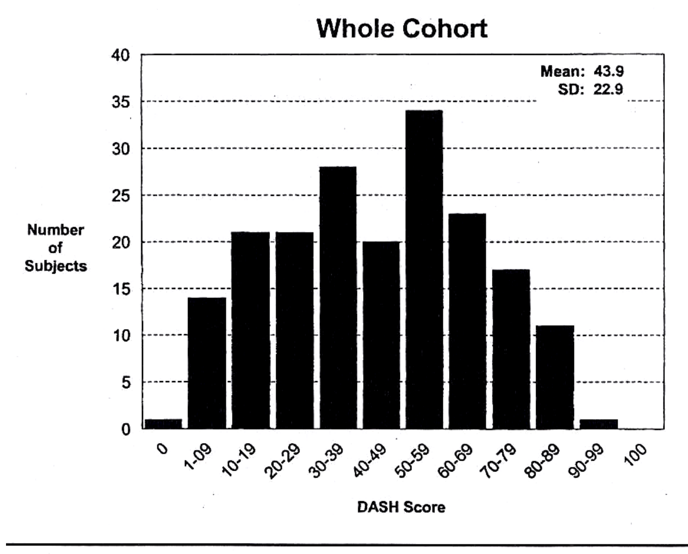
- **Validité de contenu :** elle est assurée par les avis d'experts tout au long du processus de développement du DASH.
- **Validité de critères :** elle n'a pas été testée pour le DASH en l'absence de « gold standard » connu pour le membre supérieur.
- **Validité de construction :** il existe plusieurs types de validité de construction, convergente, divergente et avec des groupes connus. Il faut répondre à la question de savoir si le DASH mesure ce qu'il est censé mesurer ?

La validité convergente étudie l'importance avec laquelle une mesure est corrélée à une autre mesure construite de manière identique. Par exemple le score DASH élevé sera corrélé avec un score de douleur élevé du membre supérieur.
La validité divergente décrit comment des mesures construites différemment ne sont pas bien corrélées. Par exemple le score DASH sera peu corrélé aux scores de douleur de la hanche.
La validité avec groupes connus implique de définir des patients appartenant à différents groupes qui sont en rapport avec la construction à mesurer. Par exemple, les patients avec des symptômes sévères du membre supérieur devraient avoir plus d'incapacité, donc des scores DASH plus élevés, que des patients avec des symptômes modérés.

Première question : le DASH est il capable de décrire des personnes avec une pathologie du membre supérieur ?

Il est important pour un outil comme le DASH de savoir s'il produit un effet « plafond » ou un effet « plancher » ce qui correspond aux scores des deux extrêmes. Si ces effets sont trop importants, c'est-à-dire si trop de patients se trouvent aux extrémités il y aura une perte de sensibilité pour détecter l'incapacité et des difficultés pour mesurer un changement lors d'une mesure itérative.

Les études (10,11) et la cohorte qui a servi à la mise au point du DASH, montrent des scores bien répartis avec une courbe en cloche et pas d'effets plafond ou plancher (cf courbe). Dans l'étude de Beaton (11) les scores DASH ont pu discriminer les pathologies d'épaule et les pathologies de la main/poignet. Les scores DASH des épaules étaient nettement supérieurs. Dans différentes études le score moyen du DASH va de 27,5 (69) à 43,9 (11). Dans aucune étude, sauf celle de SooHoo (84), il n'a été rapporté un score maximal de 100 et très rarement un score de 0 (2 ou 3 patients isolés).



Baseline distribution of DASH scores (out of 100) for the entire sample (n=200) of the Beaton

Deuxième question : existe-t-il une validité de construction convergente ?

Les corrélations attendues entre le DASH et d'autres mesures devront être modérées (environ 0.40) à hautes (>0.75) pour être significatives.

Le DASH a montré des corrélations avec :

- D'autres questionnaires du membre supérieur :
  Le Boston questionnaire (questionnaire sur le canal carpien) dans deux études (11, 44).
  Le questionnaire PRWE (patient rated wrist evaluation) dans une étude (65).
  Le POS (Patient Outcomes of Surgery han/arm), questionnaire post chirurgical (18).
  Des questionnaires sur l'épaule.
- Le COPM (canadian occupational performance measure), questionnaire général. d'auto-évaluation (20,92).
- Avec le SF 36 ou le SF 12, qui est un instrument générique très connu et très utilisé pour évaluer l'état de santé d'un patient. Il a plusieurs dimensions : physique, douleur, mental, émotionnel.
  Lors de l'étude de mise au point du DASH des corrélations ont été retrouvées avec le SF 36 fonction physique, douleur et composant physique (83).
  Jain (44) a montré des corrélations avec la fonction physique et la douleur.

Offenbacher (71) a montré des corrélations avec la douleur et à moindre degré avec le composant physique.
SooHoo (84) publie des résultats qui vont dans le même sens, à savoir une corrélation modérée de 0.58 à 0.72.
Lee (54) lors de la publication du DASH en chinois retrouve des valeurs plus faibles entre 0.42 pour le rôle physique et 0.48 pour la douleur qui sont significatives avec un $p<0.01$
Atroshi (7) a montré des corrélations avec le SF12, version plus courte du SF 36, surtout pour le composant physique. Les patients qui avaient les scores DASH les plus élevés rapportaient un mauvais état de santé général.
Padua (73) lors de la publication du DASH en italien rapporte des corrélations avec la douleur, la fonction physique mais aussi et cela est plus inhabituel une corrélation avec le composant mental, ce qui n'est pas retrouvé dans les autres études.

- Le HAQ (Health Assessment Questionnaire) qui est aussi une mesure très utilisée dans le domaine de l'incapacité physique. Offenbacher (72) a montré une corrélation forte avec un r=0.88 entre le DASH et l'HAQ.

- Différentes échelles visuelles analogiques (EVA) :
Dans l'étude de Beaton (11) des corrélations positives sont démontrées entre le score DASH et l'EVA douleur, l'EVA fonction, l'EVA appréciation de la sévérité du problème et l'EVA capacité à travailler. Toutes ces EVA étaient remplies par le patient. Donc plus le score DASH était élevé plus l'EVA l'était aussi.
Offenbacher (72) et Lee (54) rapportent une corrélation entre score DASH et douleur mesurée sur une EVA, de 0.81 et 0.49.

- Des données cliniques :
Navsarikar (69), dans une population de 50 rhumatismes psoriasiques a montré une corrélation entre la force de serrage de la main, le nombre d'articulations en poussées du membre supérieur, la classe fonctionnelle selon la classification de l'ACR(American College of Rhumatology) et le score DASH.
Westphal (95) a montré chez 109 patients ayant une fracture de Pouteau-Colles, des corrélations entre le DASH et les déficits des amplitudes articulaires du poignet.
Lee (54) rapporte une corrélation avec la force de serrage.

Le tableau ci-dessous résume la validité convergente du DASH **(Tableau n°16).**

COMPARATIF	ETUDES	RESULTATS
Boston Questionnaire (BQ)	Beaton (11)	r=0.89 (fonction) r=0.71 (symptômes)
	Jain (44)	r=0.85
PRWE	McDermid (65)	r= 0.72
POS	Cano (18)	r= 0.63 à 0.88
COPM	Veehof (92)	r=0.79
	Case-Smith (20)	r=0.55 à 0.68
SF 12	Atroshi (7)	SF12 physique r= 0.74
SF 36	Etude pilote (83)	SF36 fonction physique r=0.65 SF 36 douleur r=0.70 SF 36 score physique global r=0.73
	Jain (44)	SF 36 fonction physique r=0.68 SF 36 douleur r=0.56
	Offenbacher (71)	SF 36 douleur r= 0.79 SF 36 fonction physique r=0.58
	SooHoo (84)	SF 36 douleur r=0.43 SF 36 fonction physique r=0.59 SF 36 rôle physique r=0.61
	Lee (54)	SF 36 douleur r=0.48 SF 36 score physique global r=0.49
	Padua (73)	SF 36 mental r=0.70 SF 36 douleur r= 0.51 SF 36 score physique global r=0.41
	Gay (34)	SF 36 rôle physique r=0.46 SF 36 douleur r= 0.43
HAQ	Offenbacher (71)	r=0.88
EVA douleur	Beaton (11)	r=0.72
	Offenbacher (72)	r=0.81
	Lee (54)	r=0.49
EVA sévérité du problème	Beaton (11)	r=0.69
EVA fonction	Beaton (11)	r=0.79
EVA travail	Beaton (11)	r=0.77
Force de serrage	Navrisankar (69)	r=0.47
	Lee (54)	r= 0.43 et 0.44
Articulations en poussée du Membre sup (rhum pso)	Navrisankar (69)	r=0.65
Classe fonctionnelle ACR (rhum pso)	Navrisankar (69)	r=0.45
Déficit mobilités du poignet (fracture Pouteau-Colles)	Westphal (95)	r=0.53

Troisième question : existe-t-il une validité de construction divergente ?
Il est aussi important que le DASH ne montre pas de corrélation avec des mesures de différents concepts.
Le DASH n'est pas construit pour détecter la dimension mentale ou émotionnelle. Cela a été bien montré par la comparaison avec le SF 36 dans différentes études (10, 44) avec des corrélations inférieures à 0.40 : 0.38 avec le score de composant mental du SF 36 dans l'étude de Jain (44) et 0.29 dans l'étude de Beaton (10). Pour l'index de santé mentale du SF36 la corrélation est de 0.29 et pour la fonction émotionnelle, elle est de 0.37 dans l'étude de Beaton.
On retrouve les mêmes résultats dans les études suivantes : Offenbacher (71,72), McDermid (62,63),
A noter dans l'étude de Padua (73) déjà citée ci-dessus une corrélation positive entre le score SF 36 mental et le score global composant mental et le DASH respectivement de 0.70 et 0.55.
SooHoo (84) trouve des résultats qui vont dans le même sens que Padua sur une série de 90 patients. Avec une corrélation de 0.59 pour le composant mental global et de 0.62 pour le composant émotionnel. Ces corrélations sont modérées mais les auteurs n'apportent pas d'éléments de réponse pour les expliquer. Mais ils soutiennent tout de même la validité de construction du DASH comme instrument de mesure pour les pathologies du membre supérieur.
Globalement les études qui confirment l'absence de corrélation du DASH avec les mesures de dimension mentale ou émotionnelle l'emportent sur les études qui montrent le contraire. Il est difficile de comparer ces séries qui comportent des populations différentes, des pathologies différentes de localisations différentes. Il serait intéressant de disposer de séries homogènes par localisation et pathologie et de comparer le DASH au SF 36. On peut imaginer que les résultats sont différents selon les pathologies avec un retentissement mental et émotionnel variable.
Navrasikar (69) ne trouve pas de corrélation (r=0.09) entre le score DASH et le nombre d'articulations atteintes mais non actives du membre supérieur, dans le rhumatisme psoriasique. Il faut rappeler, dans le rhumatisme psoriasique, qu'il existe une corrélation entre le score DASH et le nombre d'articulations actives du membre supérieur (cf.supra). Les articulations actives produiraient des symptômes détectés par le DASH alors que les articulations atteintes ne sont pas nécessairement siège d'une limitation fonctionnelle gênante pour le patient.

Quatrième question : le DASH est il capable de discriminer différents groupes de patients ?
Le DASH a montré dans différentes études sa capacité à discriminer des groupes de patients. Nous allons voir ces différentes études.

Jain (44) dans son étude de 43 patients ayant été opérés d'un conflit ulno-carpien a demandé aux patients de s'auto évaluer à un an avec une échelle de 0= pas de problème à 4=très sévère. Il a comparé les résultats de cette auto évaluation aux scores DASH. Il existe une corrélation entre les scores DASH et les scores d'auto évaluation par le patient. Le DASH est à même de faire 5 groupes de patients : très sévère 59.8, sévère 46.4, modéré 39.1, moyen 12.4 et pas de problème 11.4. Dans cette étude le score DASH a pu discriminer entre ceux qui travaillent avec des restrictions en rapport avec le poignet de ceux qui travaillent sans restrictions. (Score moyen 36.4 versus 18.9).

Dans l'étude de mise au point du DASH (83), les auteurs ont pu montrer que le score DASH était corrélé à l'évaluation par le thérapeute de la gravité du problème du patient. Là aussi 5 groupes ont été mis en évidence : très sévère (55.6), sévère (49.9), modéré (37.5), léger (24.2) et très léger (17.8). Ces résultats sont assez proches de ceux de l'auto évaluation du patient, exposés ci-dessus. Dans cette étude les scores DASH reflétaient bien la perception par le thérapeute de comment le patient va. Dans cette même étude les scores DASH ont pu discriminer les patients capables de travailler (score moyen de 32) de ceux incapables de le faire (score moyen 56). Cela supporte le fait que le DASH est une mesure d'incapacité.
Beaton (11) montre les mêmes résultats pour le travail avec un score moyen de 26,8 pour ceux qui peuvent travailler et de 50,7 pour ceux qui en sont incapables. Pour les activités de la vie quotidienne le score est de 47,1 pour ceux qui ne sont pas capables et de 23,6 pour ceux qui le sont.
Krimmer (53) en 2000 dans une série de 138 patients opérés, soit d'une arthrodèse complète du poignet soit d'une arthrodèse du médio carpe pour une arthrose intra carpienne, a montré que le score DASH pouvait discriminer les patients opérés avec les deux techniques. En effet, les patients opérés par arthrodèse médio carpienne avaient des scores DASH beaucoup plus bas (score moyen de 33) par rapport à ceux opérés par arthrodèse totale (score moyen de 45) avec une différence significative ($p<0.001$). Ceci confirmait le fait que la solution de l'arthrodèse médio carpienne quand elle est possible est la meilleure solution fonctionnelle.
Westphal (95) sur 107 patients opérés d'une fracture de Pouteau-Colles a montré que le score DASH pouvait discriminer les patients avec différents types de fracture (fractures de type A, B ou C). Les fractures de type A ou B avaient un score DASH plus bas que les fractures de type C.
Jester (47,48) chez 590 patients ayant été opérés de la main pour différentes pathologies a évalué par le DASH les altérations et les incapacités ressenties par ces patients 2 à 5 ans après. Il a retenu 11 groupes de diagnostics. Il a ensuite, par une analyse statistique, étudié les relations entre ces groupes diagnostiques, à l'échelle individuelle des items du DASH. Les scores DASH ont montré qu'il existait des différences entre les groupes de pathologies. Par exemple des scores plus élevés étaient notés dans les groupes de lésions sévères (lésion du nerf médian, brûlures) par rapport aux pathologies moins graves. Dans tous les groupes, les scores concernant les symptômes étaient plus élevés que les scores concernant l'activité fonctionnelle. L'idée des auteurs est que le DASH est capable de capturer les limitations du patient sur la base d'un item individuel. Le DASH pourrait alors aider le clinicien à faire un profil spécifique des limitations engendrées par une pathologie, dans le but de mettre en route un programme de réhabilitation ciblée. Cette hypothèse reste à confirmer dans d'autres études.

En conclusion : la validité de construction du DASH est prouvée dans différentes études, nous l'avons vu en détails. Mais, comme les auteurs du DASH le disent (83), cette validité n'est jamais définitive et reste toujours à conforter par des études, qui apporteront chacune une pierre de plus à l'édifice du DASH. Un instrument n'est jamais universellement fiable ou valide. Il est fort probable que dans les années à venir, les études sur le DASH seront nombreuses avec des populations variées, des pathologies variées et que cet outil sera de plus en plus utilisé en pratique courante.

V.2.3.3 Sensibilité au changement (83).

Elle est très importante puisqu'elle prouve la capacité de l'instrument à mesurer un changement, par exemple une amélioration clinique chez un patient après un traitement (chirurgie, traitement médical, kinésithérapie, ergothérapie)

Types de changement: pour évaluer la sensibilité d'un instrument ou pour interpréter des études sur le changement, le type de changement doit être considéré avec prudence. Un outil peut être sensible avec une construction mais pas avec une autre.

- ***Le changement minimal potentiellement détectable*** : il faut regarder le nombre d'items et l'échelle de l'outil. Pour le DASH, il y a 30 items et 5 réponses possibles. Le score va de 30 à 150 points. Le changement d'une catégorie de réponse dans un item est le changement minimum possible pour chaque patient.
- ***Le changement minimal détectable (MDC)*** : c'est le changement minimum du score observé qui n'est pas attribuable à l'erreur de mesure normale du test. Le MDC est calculé en utilisant l'erreur standard de mesure (SEM) pour les changements de score et est exprimée en fonction de niveaux de confiance variables (90 ou 95 %). Beaton (10), a calculé le MDC à 90 et 95 % d'intervalle de confiance, et rapporte un MDC 90% à 10.7 points DASH et un MDC 95% à 12.75 points DASH. Cela veut dire que si un patient a un changement de 10.7 points dans les score DASH, on peut être sur à 90 % que ce patient a un changement dans l'incapacité de son membre supérieur, plutôt qu'une erreur de mesure.
- ***Le changement observé*** : c'est le changement de ce qu'on mesure par exemple avant et après un traitement, c'est-à-dire dans deux différentes occasions. Par exemple on décide de voir l'effet de la chirurgie sur le canal carpien, on pourra alors tester le patient avec le DASH avant la chirurgie puis 6 semaines après, par exemple. On fait l'hypothèse que dans la majorité des cas les patients sont améliorés, voire guéris par l'intervention ce qui est le plus souvent le cas.
- ***Le changement estimé*** : c'est le changement observé mesuré dans un groupe, identifié par un indicateur externe qui a montré une amélioration ou une détérioration. Cet indicateur peut être la réponse à une simple question : par rapport à avant le traitement, êtes vous mieux, pareil ou pire ? L'indicateur externe est ce qui différencie le changement observé du changement estimé.
- ***Le changement important*** : c'est la même chose que le changement estimé mais ce changement est important. Le même changement peut être vu comme important pour un patient et peu important pour un autre. Par exemple, après un traitement le score DASH peut être plus bas, les patients ne rapporteront pas l'amélioration de la même façon. Ce type de changement est hautement individualisé et l'interprétation de l'importance de l'amélioration est variable selon la perspective (du patient, du clinicien, des assurances et de la société).

Statistiques utilisées pour décrire la sensibilité au changement : les méthodes sont nombreuses et variées.

- ***Comparaisons des scores à To et à Tn :*** c'est la méthode la plus simple. Un changement dans les scores pourra être considéré comme une évidence de sensibilité au changement. On peut aussi rechercher des corrélations entre la diminution du score DASH et la diminution d'une autre mesure connue du membre supérieur.

- ***Calcul de l'effet de taille (size effect) et de la réponse standardisée moyenne (SRM) :*** l'effet de taille est calculé en divisant la moyenne du changement observé par l'écart type du score de base. La SRM est calculée en divisant la moyenne du changement du score par l'écart type des changements des scores. L'avantage de ces deux mesures est qu'elles peuvent être traduites en une unité standard de mesure, qui permet des comparaisons avec d'autres instruments quand ils sont appliqués chez des patients identiques pour la même application. Le désavantage est que ces statistiques sont utiles seulement au niveau d'un groupe et pas au niveau individuel. Un effet de taille et /ou une SRM entre 0.2 et 0.5 est considéré comme petit, entre 0.5 et 0.8 comme modéré et supérieur à 0.8 comme grand.
- ***La courbe ROC (receiving operator characteristic) :*** l'avantage est qu'elle donne des informations qui peuvent être utilisées au niveau individuel. Elle permet de distinguer entre ceux qui sont mieux et ceux qui ne le sont pas.

Pour interpréter ces résultats, les chiffres seuls n'ont pas de valeur. Il faut avoir des informations sur le groupe étudié : potentiel de changement, le traitement appliqué, l'intervalle entre les tests et le type de changement évalué. Cela permet d'interpréter au mieux la sensibilité au changement de l'instrument et d'utiliser ces données dans une situation clinique ou de recherche donnée.

Données des différentes études concernant la sensibilité au changement :

D'assez nombreuses études se sont intéressées à ce problème. Nous allons les étudier en reprenant la classification utilisée plus haut pour les types de changement.

- *Changement observé : sensibilité au changement du DASH avant et après traitement quand le changement était attendu.*
  Beaton (11) dans cette grande étude concernant 172 patients ayant des pathologies variées du membre supérieur, a montré que le DASH était sensible au changement observé 12 semaines après un traitement chirurgical ou non. La majorité des patients avaient une amélioration du score DASH ce qui indique une diminution de l'incapacité du membre supérieur. Les SRM étaient entre 0.74 et 0.81. Cette étude a aussi décrit la sensibilité du DASH, en corrélant les changements de son score avec les changements rapportés de l'intensité de la douleur, de la fonction et de la sévérité du problème mesurés avant le traitement et 12 semaines après. Les r étaient tous entre 0.65 et 0.69.
  McDermid (62), a évalué la sensibilité au changement de trois questionnaires le DASH, le PRWE et le SF36 après fracture du poignet. 59 patients ont été évalués au début puis à 3 et 6 mois. Le DASH était plus sensible que le SF 36 mais moins que le PRWE, ce qui est logique puisque le PRWE est un questionnaire spécifique du poignet. Malheureusement il ne rapporte pas les valeurs moyennes du DASH. Les SRM sont de 2.01 à 3 mois et 2.52 à 6 mois.

- *Changement estimé : sensibilité du DASH chez ceux qui ont estimé s'être amélioré sur une échelle auto évaluée.*
  Beaton (11) a demandé aux patients, après 12 semaines de traitement, si leur problème de membre supérieur avait changé et de combien. L'échelle comportait 11 points et au-delà de 6 le patient était considéré comme amélioré. Les patients étaient aussi interrogés avec la même échelle sur l'amélioration de la fonction. Dans les deux cas le DASH était sensible avec des SRM de 1.06 et 1.20. Une analyse avec les courbes ROC a été pratiquée en utilisant les réponses à 12 semaines à la question suivante :

faites vous face avec votre problème et faites vous ce que vous auriez envie de faire ? (Étude du coping=faire face). Une réponse OUI a été le critère utilisé pour l'amélioration et a été comparé aux changements des scores DASH. Un changement de 15 points a permis de distinguer les patients arrivant à faire face, de ceux n'y arrivant pas (sensibilité 0.67, spécificité 0.78 précision 68%).
Gummesson (39), en 2003, a administré le DASH à 109 patients opérés de diverses régions du membre supérieur, en pré opératoire et 6 à 21 mois après l'intervention. La variation du score a été analysée par l'effet de taille et les SRM. Comme Beaton, ces résultats ont été analysés aussi en fonction d'une échelle d'auto évaluation du résultat par le patient. La moyenne des scores DASH étaient de 35 en pré opératoire et de 24 en post opératoire et la moyenne de changement du score était de 15. L'effet de taille était de 0.7 et le SRM de 1,2. Le changement moyen du score DASH (intervalle de confiance 95%) pour les patients rapportant que leur bras était beaucoup mieux ou beaucoup moins bien après chirurgie était de 19, et pour ceux qui se disaient un peu mieux ou un peu moins bien était de 10 (p=0.01). Si on prenait la chirurgie du canal carpien l'effet de taille était de 0.7 et la SRM de 1.0. Pour les auteurs, une différence de 10 points dans la moyenne des scores DASH peut être considérée comme le changement minimal important.

- *Changement important : sensibilité du DASH chez les patients rapportant un changement important.*
  Beaton (83) a interrogé les patients de sa cohorte et leur a demandé si l'ensemble de leurs problèmes (douleur, fonction, capacité à travailler) avaient changé, et si ce changement était important pour eux. Les sujets avec un changement positif et important étaient considérés comme ayant une amélioration importante. Les SRM pour cet important changement étaient de 1.06 (douleur) à 1.27 (fonction).

- *Comparaisons avec des mesures spécifiques pour une articulation.*
  McDermid (62) a montré que le questionnaire PRWE était plus sensible que le DASH à 3 et 6 mois après une fracture du poignet.
  Beaton (11) a montré que le DASH avait une meilleure sensibilité que le Boston questionnaire et que le SPADI (questionnaire sur l'épaule) pour la population globale des patients et pour les sous-groupes en fonction de la région du traumatisme (épaule et poignet/main).
  McDermid (65) dans une deuxième étude a comparé le DASH au PRWE modifié dans une cohorte de 60 patients (36 mains et 24 poignets). Ces questionnaires ont été administrés à 0 et 3 mois. L'effet de taille et les SRM ont été calculés. La sensibilité des deux questionnaires était similaire avec un effet de taille supérieur à 0.80. Il y avait une tendance pour le PRWE à être plus sensible que le DASH (SRM de 1.51 versus 1.37), mais cette différence était minime. Il y avait une forte relation entre le changement rapporté par le DASH et le PRWE (r=0.72).
  Gay (34), en 2003, a comparé dans une étude sur 34 patients ayant un canal carpien, le DASH, le Boston questionnaire (BQ) spécifique du canal carpien et le SF36. Les patients remplissaient les questionnaires en pré opératoire et 6 et 12 semaines après. L'instrument le plus sensible à 12 semaines était le BQ (effet de taille 1.71/SRM 1.66) suivi du DASH (1.01/1.13) puis du SF36. Il y avait une bonne corrélation entre les changements de scores au DASH et au BQ (coefficient de corrélation de Spearman à 0.87). Les auteurs concluent que le DASH est suffisamment sensible au changement pour être utilisé comme une mesure d'outcome après chirurgie du canal carpien.

Greenslade (38) en 2004, a fait la même étude que Gay sur 57 patients opérés d'un canal carpien. Les questionnaires sont administrés en pré opératoire et 3 mois après l'intervention. Les résultats sont les mêmes que dans l'étude Gay. La moyenne d'amélioration du score DASH à 3 mois post opératoire était de 15. Il y avait une corrélation entre les scores pré opératoires et les changements du score, plus le score était haut en pré opératoire plus l'amélioration du score était importante. La SRM du DASH était de 0.66, celle du BQ fonction de 0.62 et celle du BQ sévérité des symptômes de 1.07. Les conclusions des auteurs rejoignent celles de Gay.
Case-Smith (20), a comparé le DASH au COPM (canadian occupational performance measure) et au SF 36. 33 patients ont eu en moyenne 13 heures d'ergothérapie, puis ont été évalués par le DASH, le SF 36 et le COPM au début et à la fin du traitement. Les scores du DASH à T0 étaient en moyenne de 62.8 et après traitement de 36.2 soit une amélioration en moyenne de 26.6 points. L'effet de taille était de 1.50. Ceci était comparable à l'amélioration du score du COPM. Les corrélations entre DASH et COPM étaient modérées entre 0.55 et 0.68. Le COPM d'après les auteurs était le plus sensible aux changements du patient, suivi par le DASH et le SF 36.
Cano (18) a comparé le DASH au questionnaire POS (Patients Outcomes of Surgery Hand/Arm cf chapitre suivant) et au MHQ (Michigan Hand Questionnaire). Les corrélations entre le DASH et le POS sont bonnes entre 0.63 et 0.88.
Kotsis en 2005 (52) a comparé le DASH et le MHQ chez 50 patients après chirurgie du canal carpien. Il s'agit d'une étude prospective avec administration des questionnaires avant et 6 mois après l'intervention. Le DASH est sensible au changement avec une SRM moyenne de 0.7, équivalente à la SRM du MHQ fonction (0.6). La SRM du MHQ douleur est à 0.9.
Tout ceci montre globalement que le DASH a une sensibilité comparable aux questionnaires spécifiques d'une articulation et que cet outil est utile pour mesurer un changement pour des patients avec des pathologies variées du membre supérieur.
Le tableau ci-dessous résume les différentes études de la sensibilité au changement.

TYPE DE CHANGEMENT	ETUDES	COLLECTE DONNEES	SCORES DASH B : base S : suivi C : moyenne chgt(DS)	SRM
Changement observé	Beaton (11)	T 0 et 3 mois	B : 44.5 (22.7) S : 30.9 (22.8) C : – 13.3 (16.9)	Total 0.78 Epaule 0.81 Main poignet 0.74
	McDermid (62)	T0, 3 et 6 mois	Non reportés	3 mois : 2.01 6 mois 2.52
Changement estimé	Beaton (11)	T 0 et 3 mois	B : 42.9 (22.9) S : 24.9 (20.2) C : - 17.3 (16.4) Pour ceux qui disent que le problème va mieux	Total 1.06 Epaule 1.13 Main/poignet 0.92
			B : 40.7 (23.4) S : 20.2 (19.2) C : - 19.7 (16.5) Pour ceux qui estime leur fonction meilleure	Total 1.20 Epaule 1.44 Main/poignet 0.91
			Changement du score pour déterminer ceux qui sont capables de faire face de ceux qui ne sont pas capables	Changement du score le plus précis:15
	Gummesson (39)	Pré opératoire et 6 à 21 mois après	B : 35 S : 24 C : +/- 15. +/-19 pour ceux qui disent être beaucoup mieux ou beaucoup moins bien. +/- 10 pour ceux qui disent être un peu mieux ou un peu moins bien	Effet de taille 0.7 SRM 1.2 Changement minimal important 10 points pour le score DASH
Changement important	Beaton (83)		Estimation du changement par le patient et si ce changement est important	1.06 douleur 1.27 fonction

Comparaisons avec mesures spécifiques	McDermid (62) PRWE	T 0 3 et 6 mois	Non rapportées	3 mois 2.01 6 mois 2.52
	McDermid (65) PRWE	T 0 et 3 mois	?	Effet de taille >0.80 SRM PRWE : 1.51 SRM DASH ; 1.37 r= 0.72 entre les 2 questionnaires
	Gay (34) Boston Questionnaire (BQ) et SF 36	T 0 puis 6 et 12 semaines	?	BQ : effet de taille 1.71 SRM 1.66 DASH : effet de taille 1.01 SRM 1.13 R=0.87 entre les 2 questionnaires
	Greenslade (38) BQ et SF 36	T 0 et 3 mois	?	DASH SRM : 0.66 BQ SRM 0.62 et 1.07
	Case-Smith (20) COPM SF 36	T0 et 6 à 8 semaines après	B : 62.8 S : 36.2 C : 26.6	Effet de taille DASH 1.50 COPM 2.45 et 2.52. SRM ? Corrélations entre les 2 questionnaires 0.55 à 0.68
	Cano (18) POS et SF 36	T0 et à 3 mois	?	Corrélation entre DASH et POS 0.63 à 0.88 SRM ??
	Kotsis (55) MHQ	T0 et 6 mois	?	SRM DASH 0.6 SRM MHQ fonction 0.7 SRM MHQ douleur 0.9

Tableau n°17 : sensibilité au changement du DASH.

En conclusion : nous avons vu que la sensibilité au changement du DASH a été bien étudiée et qu'elle est complexe puisqu'elle englobe plusieurs concepts. Comme la fiabilité et la validité, la sensibilité au changement sera toujours évolutive et de nouvelles informations seront apportées par les études publiées dans le futur. Ceci renforce le fait que cet outil est utile et a été développé de façon optimale et devrait être utilisé plus largement.

V.2.3.4 Interprétation du score DASH à l'échelle du patient. Qu'est ce que le score veut dire ?

Même si les propriétés psychométriques du DASH sont bonnes, la question de savoir si le DASH est utile à l'échelon du patient doit être posée. Toutes les mesures ont une marge d'erreur y compris les questionnaires. Différents critères sont indispensables pour utiliser un outil d'évaluation à un échelon individuel.

- *L'instrument :* le premier critère est le caractère pratique et facile d'utilisation avec le patient. Il faut que le patient comprenne les questions et sache lire. Les versions traduites existantes peuvent permettre de l'administrer facilement à un patient ayant une langue maternelle différente de la langue habituellement utilisée dans le pays. Il est facile à administrer avec un temps moyen de 10 minutes environ. Il peut être administré dans la salle d'attente avant une consultation, dans la chambre d'hôpital ou bien être envoyé par courrier pour être rempli à domicile. Le DASH a toutes ces qualités. Nous avons vu au chapitre précédent, qu'un petit nombre de patients avaient des scores soit très élevés soit très bas. Le DASH est donc capable de capturer la plus grande partie des symptômes et de la fonction physique des patients ayant un problème de membre supérieur.

- *La précision de la mesure :* c'est le deuxième critère à remplir. Il faut estimer l'erreur autour du score DASH observé, aussi bien à un moment donné (précision transversale) qu'au cours du temps (précision longitudinale).
  La précision transversale fournit une étendue de scores à l'intérieur desquels le vrai score existe. Cette variation, nous l'avons déjà vu, s'appelle l'erreur standard de mesure (SEM). Pour la calculer on utilise le coefficient de fiabilité Cronbach alpha. Ensuite on définit l'intervalle de confiance en général 95 %. Dans l'étude princeps du DASH (83) la valeur du SEM était de 4,4, ce qui donnait des variations de +/- 8,8 points pour un intervalle de confiance à 95%. Pour un score de 50, on peut être sur à 95 % que le vrai score se trouve entre 41 et 59.
  La précision longitudinale s'intéresse à la mesure au cours du temps, quand on doit la répéter. Dans ce cas, l'erreur associée au changement du score doit inclure l'erreur de mesure à T0 et à Tn. C'est là qu'intervient la MDC, déjà évoquée plus haut. C'est le minimum de changement du score qui n'est pas attribuable à l'erreur normale du test. Son calcul est différent de celui de la précision transversale, elle fait appel au coefficient de corrélation test retest. On choisit là aussi un intervalle de confiance de 95%. Beaton (11) a calculé une MDC95% à 12.7 points dans son étude. En d'autres termes, si le score du patient a changé de plus de 13 points, il y a moins de 5% de chances que le patient n'ait pas vraiment changé. On peut remarquer que si le coefficient de corrélation test retest est assez bas la MDC sera plus importante et il sera plus difficile de dire si le patient a changé étant donné la grosse marge d'erreur, d'où l'importance d'avoir un coefficient de corrélation test retest le plus haut possible.

- *Validité du score à l'échelon individuel* : nous avons vu dans le chapitre précédent que la majorité des études de validité et de sensibilité au changement étaient faites à l'échelon du groupe. Il faudrait pouvoir interpréter cela au niveau de l'individu pour qu'il y ait une pertinence clinique.
  La validité transversale à l'échelon individuel est encore difficile à apprécier car le DASH est un outil récent et que des normes strictes ne sont pas encore publiées. On ne

peut pas dire encore quels scores DASH constituent une incapacité moyenne, modérée ou sévère et si un individu est capable de travailler ou non. Dans des études non publiées et dans l'étude princeps, les scores des patients capables de travailler étaient de 20 à 30, alors que ceux qui étaient incapables de travailler étaient vers 50 à 60. Des normes du DASH ont été collectées chez 1800 personnes n'ayant pas de pathologie du membre supérieur. Les résultats sont disponibles par tranches d'âge et par sexe (83) et sont disponibles sur les sites internet de l'AAOS : www.aaos.org ou du DASH www.dash.iwh.on.ca.
La validité longitudinale à l'échelon individuel est étudiée par la MCID (Minimal Clinically Important Difference). En fait, cette mesure ne serait pas aussi facile à utiliser qu'on le dit. Elle peut être calculée selon différentes méthodes et différents points de vue. Il serait plus utile d'étudier le changement important du point de vue du patient. Beaton (10,11) nous l'avons déjà vu a suggéré qu'un changement de 15 points dans le score DASH, est le meilleur moyen de séparer les patients qui s'améliorent de ceux qui ne s'améliorent pas. La MDC95 serait en fait a peu près équivalente à la MCID, il semble plus facile d'utiliser cette valeur tant que des travaux plus importants n'auront pas été publiés sur la MCID.

V.2.4 Données diverses.

V.2.4.1 Versions traduites.

Le DASH est protégé par copyright et toute traduction du DASH doit faire l'objet d'une procédure stricte. Les versions validées officiellement sont les suivantes : arménienne, chinoises (2 versions de Hong-Kong, 1 version de Taiwan), canadienne française, française, allemande, grecque, hébreu, japonaise, norvégienne, russe, espagnoles (1 Espagne, 1 Porto Rico), suédoise et turque. La version italienne (73) a été publiée et a fait l'objet d'une traduction dans les règles mais elle n'est pas validée officiellement par les auteurs du DASH. A noter qu'il existe deux versions allemandes (36,71) mais seule la version de Germann (36) est validée officiellement.
Pour valider la traduction, il doit y avoir 5 traductions de l'anglais à la langue donnée. Il en résulte une version dans la langue donnée. Puis deux contre traducteurs traduisent à nouveau de la langue donnée à l'anglais cette version. Puis ces traductions en anglais sont comparées à la version source. Puis la version traduite doit être administrée avec aide à des patients pour s'assurer que la compréhension soit bonne, puis sans aide pour évaluer le remplissage et le taux de réponses. Enfin, la version peut être validée officiellement.
Ceci permet de s'assurer du sérieux des traductions proposées et permet de garantir une fidélité à la version originale et donc de pouvoir comparer des résultats d'études publiées avec des DASH dans d'autres langues. Nous avons vu plus haut, que ces versions ont d'excellentes propriétés psychométriques (tableau n°13).

V.2.4.2 Le DASH version courte ou Quick DASH.

C'est une version courte, 11 items du DASH et toujours les 2 modules optionnels. Il n'a pas été encore publié officiellement mais il est disponible sur le site Internet du DASH. Il a globalement les mêmes propriétés psychométriques que le DASH : Cronbach alpha 0.94, fiabilité test retest avec ICC=0.94. La validité et la sensibilité au changement sont similaires. Le patient doit répondre à au moins 10 items sur 11 pour que le questionnaire soit interprétable. Seule une version norvégienne existe déjà.

V.2.4.3 Mode d'emploi de la version française.

Elle a été publiée par Dubert en 2001 (28) et est validée par les auteurs du DASH. Les explications sont tirées de l'article de Dubert.
Le questionnaire comprend :

- 30 questions :

    - ☞ 21 évaluent la difficulté à réaliser des activités précises de la vie quotidienne : 8 activités bi manuelles, 5 activités concernant la main dominante et 8 activités réalisées aussi bien par la main dominante que par la main non dominante.
    - ☞ 9 diverses : 3 sur les relations sociales, 6 sur des symptômes particuliers (douleur : 3, force : 1, mobilité : 1, sommeil : 1).

- 2 modules de 4 questions chacun optionnels (1 sur le sport et la musique, 1 sur la capacité de travail).

Technique de remplissage du questionnaire :

Pour chaque question, le patient doit choisir une seule réponse de 1 à 5 qui correspond à une incapacité croissante d'accomplir certaines activités au cours des sept derniers jours. Si le patient n'a pas eu l'occasion de pratiquer certaine de ces activités au cours des sept derniers jours, il doit entourer la réponse qui lui semble correspondre au degré de difficulté qu'il aurait eu s'il avait dû effectuer cette tâche.
Il est nécessaire de préciser au patient qu'il lui faut répondre à toutes les questions et n'entourer qu'une seule réponse par question, en insistant sur le fait que son choix doit se baser sur ses capacités globales au cours des sept derniers jours.
Si un patient demande des explications, il faut l'encourager à relire la question et à faire son maximum pour y répondre par lui-même. La question peut aussi être lue à haute voix si besoin.
Si le patient n'arrive toujours pas à répondre, il faut éviter de lui donner des explications complémentaires, car cela risque de l'influencer dans son choix. Il faut lui suggérer de passer à la suite pour éventuellement y revenir plus tard, ou de laisser la question sans réponse.
A la fin, il est recommandé de vérifier que toutes les questions soient bien remplies. Si ce n'est pas le cas, il faut s'assurer auprès du patient que cette omission est bien volontaire.
Les remerciements et les encouragements sont utiles pour créer une atmosphère de coopération.

Interprétation du questionnaire, Calcul du score :
L'addition des valeurs choisies pour les 30 questions permet d'obtenir un score total qui va de 30 (aucune gêne fonctionnelle) à 150 (gêne maximale).
Une règle de trois :

$$\frac{\text{score total (ST)} - 30}{1{,}2}$$

permet ensuite de ramener le score à une échelle de 0 à 100, plus facile à interpréter.

Si un (ou les deux) module(s) optionnel(s) sport/musique est (sont) rempli(s), le total pour chacun de ces modules peut également être rapporté à une échelle de 0 à 100 par la formule :

$$\frac{\text{score total (ST)} - 4}{0{,}16}$$

S’il y a moins de 4 réponses manquantes sur les 30 posées, le questionnaire est interprétable. L’évaluateur doit donner une valeur estimée aux réponses manquantes. Cette valeur est calculée en faisant la moyenne des valeurs remplies. Par exemple, si deux questions sont restées sans réponse, et que la moyenne des réponses remplies est de 3, le chiffre 3 sera attribué aux deux réponses manquantes avant de calculer le score total(ST) .

Si plus de quatre réponses sont manquantes ou si une seule des questions posées dans les modules sport/musique et travail est manquante le questionnaire n’est plus interprétable.

En cas de réponses multiples pour une question, la conduite à tenir est la même que s’il n’y a aucune réponse à la question, c’est à dire que la valeur attribuée ne sera pas l’une des deux réponses entourées, mais celle qui correspond à la moyenne des réponses données.

On peut aisément informatiser le formulaire de réponse qui peut être scanné avec un calcul automatique du score.

QUESTIONNAIRE :

Ce questionnaire s'intéresse à ce que vous ressentez et à vos possibilités d'accomplir certaines activités. Veuillez répondre à **toutes** les questions en considérant vos possibilités **au cours des 7 derniers jours**. Si vous n'avez pas eu l'occasion de pratiquer certaines de ces activités au cours des 7 derniers jours, veuillez entourer la réponse qui vous semble la plus exacte si vous aviez dû faire cette tâche. Le côté n'a pas d'importance. Veuillez répondre en fonction du résultat final, sans tenir compte de la façon dont vous y arrivez.

Veuillez évaluer votre capacité à réaliser les activités suivantes **au cours des 7 derniers jours**. (Entourez une seule réponse par ligne.)

		Aucune difficulté	Difficulté légère	Difficulté moyenne	Difficulté importante	Impossible
1.	Dévisser un couvercle serré ou neuf	1	2	3	4	5
2.	Ecrire	1	2	3	4	5
3.	Tourner une clé dans une serrure	1	2	3	4	5
4.	Préparer un repas	1	2	3	4	5
5.	Ouvrir un portail ou une lourde porte en la poussant	1	2	3	4	5
6.	Placer un objet sur une étagère au-dessus de votre tête	1	2	3	4	5

7.	Effectuer des tâches ménagères lourdes (nettoyage des sols ou des murs)	1	2	3	4	5
8.	Jardiner, s'occuper des plantes (fleurs et arbustes)	1	2	3	4	5
9.	Faire un lit	1	2	3	4	5
10.	Porter des sacs de provisions ou une mallette	1	2	3	4	5
11.	Porter un objet lourd (supérieur à 5 Kg)	1	2	3	4	5
12.	Changer une ampoule en hauteur	1	2	3	4	5
13.	se laver ou se sécher les cheveux	1	2	3	4	5
14.	Se laver le dos	1	2	3	4	5
15.	Enfiler un pull-over	1	2	3	4	5
16.	Couper la nourriture avec un couteau	1	2	3	4	5
17.	Activités de loisir sans gros effort (jouer aux cartes, tricoter, etc.)	1	2	3	4	5
18.	Activités de loisir nécessitant une certaine force ou avec des chocs au niveau de l'épaule du bras ou de la main. (bricolage, tennis, golf, etc.)	1	2	3	4	5
19.	Activités de loisir nécessitant toute la liberté de mouvement (badminton, lancer de balle, pèche, Frisbee, etc.)	1	2	3	4	5
20.	Déplacements (transports)	1	2	3	4	5
21.	Vie sexuelle	1	2	3	4	5

22. Pendant les 7 derniers jours, à quel point votre épaule, votre bras ou votre main a-t-elle gêné dans vos relations avec votre famille, vos amis ou vos voisins ? (Entourez une seule réponse)

1 Pas du tout 2 légèrement 3 moyennement 4 beaucoup 5 extrêmement

23. Avez-vous été limité dans votre travail ou une de vos activités quotidiennes habituelles du fait (en raison, par) de problèmes à votre épaule, votre bras ou votre main? (Entourez une seule réponse)

1 Pas du tout limité 2 Légèrement limité 3 moyennement limité 4 Très limité 5 Incapable

Veuillez évaluer la sévérité des symptômes suivants **durant** **<u>les 7 derniers jours.</u>** (Entourez une réponse sur chacune des lignes)

		Aucune	légère	moyenne	important	extrême
24.	Douleur de l'épaule, du bras ou de la main	1	2	3	4	5
25.	Douleur de l'épaule, du bras ou de la main en pratiquant une activité particulière Précisez cette activité : …………………..	1	2	3	4	5
26.	Picotements ou fourmillements douloureux de l'épaule, du bras ou de la main	1	2	3	4	5
27.	Faiblesse du bras, de l'épaule ou de la main	1	2	3	4	5
28.	Raideur du bras, de l'épaule ou de la main	1	2	3	4	5

29. Pendant les 7 derniers jours, votre sommeil a-t-il été perturbé par une douleur de votre épaule, de votre bras ou de votre main ? (Entourez une seule réponse)

1 pas du tout 2 un peu 3 moyennement 4 très perturbé 5 Insomnie complète

30. "Je me sens moins capable, moins confiant ou moins utile à cause du problème de mon épaule, de mon bras, ou de ma main"

1 Pas d'accord du tout 2 Pas d'accord 3 ni d'accord ni pas d'accord 4 d'accord 5 tout à fait d'accord

Les questions suivantes concernent la gêne occasionnée par votre épaule, votre bras ou votre main **lorsque vous jouez d'un instrument ou que vous pratiquez un sport ou les deux**. Si vous pratiquez plusieurs sports ou plusieurs instruments (ou les deux), vous êtes priés de répondre en fonction de l'activité qui est la plus importante pour vous.

Indiquez le sport ou l'instrument qui est le plus important pour vous :
____

Entourez 1 seule réponse par ligne, considérant vos possibilités durant les 7 derniers jours.

Avez-vous eu des difficultés :

	Aucune difficulté	Difficulté légère	Difficulté moyenne	Difficulté importante	Impossible
Pour pratiquer votre sport ou jouer de votre instrument **avec votre technique habituelle**	1	2	3	4	5
Pour pratiquer votre sport ou jouer de votre instrument **à cause des douleurs** de votre épaule, de votre bras ou de votre main	1	2	3	4	5
Pour pratiquer votre sport ou jouer de votre instrument **aussi bien que vous le souhaitez**	1	2	3	4	5
Pour passer le temps habituel à pratiquer votre sport ou jouer de votre instrument	1	2	3	4	5

Les questions suivantes concernent la gêne occasionnée par votre épaule, votre bras ou votre main au cours de votre travail.

Entourez la réponse qui, sur chacune des lignes, décrit le plus précisément vos possibilités durant les 7 derniers jours.

Si vous n'avez pas pu travailler pendant cette période, considérez comme "impossible" les quatre propositions suivantes.

Avez-vous eu des difficultés :

	Aucune difficulté	Difficulté légère	Difficulté moyenne	Difficulté importante	Impossible
Pour travailler en utilisant votre technique habituelle	1	2	3	4	5
Pour travailler comme d'habitude à cause de la douleur de votre épaule, de votre bras ou de votre main	1	2	3	4	5
Pour travailler aussi bien que vous le souhaitez	1	2	3	4	5
Pour passer le temps habituellement consacré à votre travail	1	2	3	4	5

V.2.5 Conclusions.

Nous avons étudié très longuement le questionnaire DASH. Ce questionnaire a maintenant presque 10 ans d'existence et est utilisé dans le monde entier. Sa mise au point est un exemple de rigueur et nous avons vu qu'il a, toutes les qualités psychométriques requises pour pouvoir être utilisé largement et avec sûreté. Il est facile à utiliser et est un bon indicateur de l'incapacité et des symptômes de patients avec des pathologies du membre supérieur, quelle que soit la région concernée. Il est capable de décrire des changements après traitement. Nous avons aussi souligné les questions encore en suspens en particulier dans l'interprétation des résultats à l'échelon individuel. Il faut garder à l'esprit comme le dit très bien Beaton (10) qu'un questionnaire tel que le DASH n'est jamais figé et universellement valide ou sensible au changement, mais que ses propriétés sont évolutives en fonction des pathologies concernées et des régions anatomiques touchées. Chaque nouvelle étude utilisant le DASH apportera sa pierre à l'édifice et contribuera à améliorer l'utilisation de cet outil qui deviendra j'en suis sûr, un outil d'évaluation du membre supérieur incontournable en association avec l'examen clinique et d'autres tests. Le tableau ci-dessous résume les caractéristiques du DASH.

PROPRIETES	DASH
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Questionnaire d'auto évaluation
4) Interprétation des résultats	A parfaire
5) Spécificité	Pathologies variées membre supérieur
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	anglais
Coût	gratuit (achat du manuel 50 $)
7) Temps demandé	
Administration	10 à 15 minutes
Entraînement du testeur	non
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	variées
Application dans une autre culture	Oui
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	+/-
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	Pas besoin
Intra observateur	Pas besoin
Test-retest	Oui excellente
Cohérence interne	Oui excellente
3) Validité	
Contenu	Oui
Construction	Oui
Relation avec AVQ	Oui
Relation avec statut physique	Oui
Distinction des groupes	Oui
Critères	Non testée
Sensibilité au changement	Oui
4) Normes	Oui pour les sujets sains en construction pour les autres

Tableau n°18: caractéristiques du DASH.

V.3 Le questionnaire UEFS (Upper Extremity Function Scale).

Pranski et collaborateurs ont publié cette échelle en 1997 (75). Elle a été développée dans une population spécifique de pathologies du membre supérieur en relation avec le travail. Le but de l'étude était de développer l'UEFS pour mesurer l'impact d'une pathologie du membre supérieur sur la capacité à faire des tâches physiques. Cette mesure devait être rapide, applicable à une variété de pathologies du membre supérieur, facile à comprendre et auto administrée.

- **Patients et méthodes :**

Pour mettre au point l'échelle, 12 questions ont été sélectionnées lors de discussion entre kinésithérapeutes, ergothérapeutes et patients. Sur les 12 items, 8 ont été sélectionnés car ils représentaient des activités faites par plus de 80 % des patients dans une étude pilote initiale de 35 travailleurs ayant une pathologie en rapport avec le travail. Les patients doivent coter leur difficulté à faire l'activité de 0= pas de problème à 10 = problème majeur pour faire l'activité. Les 8 activités choisies sont les suivantes :

1. Dormir.
2. Ecrire.
3. Ouvrir un récipient.
4. Prendre de petits objets avec les doigts.
5. Conduire une voiture plus de 30 minutes.
6. Ouvrir une porte.
7. Prendre un pot de lait dans le réfrigérateur.
8. Laver la vaisselle.

Le score est la somme des réponses et on peut tolérer une réponse manquante mais pas plus. Deux groupes de patients ont été étudiés.

Tout d'abord un groupe de 124 patients ayant une lésion du membre supérieur en rapport avec le travail et qui recevaient tous une indemnité (worker's compensation aux USA). 44 % de ces patients ne travaillaient pas. Les pathologies représentées étaient : 14% de canaux carpiens, 11% de tendinites, 47 % de douleurs chroniques du membre supérieur et 17% d'atteintes neurologiques diverses. 108 patients ont pu être évalués avec un questionnaire complet (informations démographiques, durée des symptômes, EVA douleur, échelle d'évaluation de la dépression, de l'anxiété, de la peur des symptômes et travail), un examen clinique et le questionnaire UEFS. 73 de ces patients ont pu être revus 12 à 24 mois après avec la même évaluation plus le questionnaire AIMS (Arthritis Impact Measurement Scale) qui est un questionnaire de la fonction dans les AVQ et qui contient trois sous échelles spécifiques de la main et du bras. Cette mesure est surtout utilisée dans la PR et est bien validée.

Le deuxième groupe était constitué de 216 patients avec un canal carpien qui ont participé à une étude prospective. Ils répondaient au même questionnaire que les patients du premier groupe et devaient évaluer leur douleur maximale et coter leurs symptômes de canal carpien sur une échelle de 0 à 5. Des mesures de la force (grip, pince) ainsi qu'une mesure de la conduction nerveuse étaient pratiquées. 144 patients ont pu être évalués au final et 91 ont été réévalués 18 mois après, avec la même évaluation plus le questionnaire AIMS.

Les corrélations ont été mesurées par le coefficient de corrélation de Pearson. Le t test a été utilisé pour explorer les différences entre groupes dans les scores UEFS. Les SRM ont été calculées.

Il n'y avait pas de différences marquées entre les deux groupes sauf que dans le groupe canal carpien les patients étaient plus nombreux à travailler (89% vs 47%).

- **Résultats :**

L'UEFS n'a pas montré d'effet plancher dans les deux groupes lors du premier questionnaire. Pour le suivi l'effet plancher a été atteint plus souvent dans le groupe canal carpien ce qui peut s'expliquer par l'amélioration obtenue après chirurgie.
La fiabilité test retest n'a pu être calculée dans cette étude.
La cohérence interne par le Cronbach alpha est bonne de 0.83 à 0.93 et comparable à celle de l'AIMS (0.93). Ceci confirme que les items de l'UEFS mesurent la même chose. Un seul facteur principal explique entre 42 et 56 % de la variance inter item et il n'existe pas d'autres facteurs associés ayant une telle valeur.
Pour la validité, le t test a été utilisé pour montrer que les scores UEFS à la première évaluation, pouvaient discriminer différents niveaux de sévérité dans les deux groupes, niveaux définis par le test de Phalen, la durée des symptômes et l'impact de la maladie sur le travail. Il existe aussi des corrélations modérées mais significatives entre les scores UEFS et des mesures diverses : EVA douleur (r=0.67), peur de la douleur (r=0.44), force (r= -0.40 grip et – 0.39 pince).
Il existe une corrélation entre l'UEFS et l'AIMS, r=0.81 pour les deux groupes (p<0.0001). Les deux mesures sont donc similaires dans leur capacité à mesurer les impacts fonctionnels sur le membre supérieur.
Il existe une forte corrélation entre les changements de l'UEFS et les changements dans les mesures de sévérité. Pour le premier groupe r=0.58 (p<0.0001) pour la douleur. Pour le deuxième groupe r=0.59 (p<0.0001) pour la douleur maximale, r=0.52 (p<0.0001) pour la sévérité des symptômes, r= -0.50 (p<0.0001) pour la force du grip.
La sensibilité au changement a été évaluée par les SRM. Dans le premier groupe, un sous groupe de 16 patients qui reportaient qu'ils allaient mieux a été sélectionné. Dans le deuxième groupe 55 patients qui ont été opérés du canal carpien ont aussi été sélectionnés. Dans les deux groupes, l'UEFS a démontré une sensibilité similaire ou meilleure aux autres mesures en particulier les mesures physiques telles que le test de Phalen ou la mesure de la force.

- **Discussion :**

L'UEFS a montré certaines qualités. Il est facile à administrer dans un temps court (moins de cinq minutes), il a une bonne cohérence interne et une validité partielle. Sa sensibilité au changement suggère que l'UEFS est plus à même de détecter des changements significatifs, comme le résultat de la progression d'une maladie, d'un traitement ou d'une intervention préventive, que les mesures traditionnelles. La forte corrélation avec le travail suggère que l'UEFS mesure des fonctions qui sont importantes pour la capacité à travailler, même chez des patients qui sont pensionnés pour leurs problèmes de membre supérieur. Ceci suggère une association entre la fonction à la maison et au travail, car l'UEFS n'explore que des activités non professionnelles. L'UEFS n'explore que la dimension physique et pour explorer les autres dimensions il vaut mieux utiliser le SF 36 par exemple.
Dans cette étude il existe des limitations. La fiabilité test retest n'a pas été étudiée. Il n'y a pas eu de groupes qui disaient aller moins bien avec le temps, ce qui aurait permis d'étudier la sensibilité à la détérioration. La corrélation entre l'amélioration du score et l'amélioration de la fonction n'est pas prouvée même pour le groupe canal carpien. Une des limites de ce questionnaire est son utilisation dans un groupe de patients restreints de pathologies du membre supérieur en rapport avec le travail et il est assez curieux qu'aucune des questions n'ait un rapport avec l'activité professionnelle. Après sa publication ce test ne semble pas avoir fait l'objet d'autres publications ce qui fait qu'on ne peut le retenir actuellement pour une utilisation en routine.
Le tableau ci-dessous résume les propriétés de ce questionnaire.

PROPRIETES	UEFS
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Questionnaire d'auto évaluation
4) Interprétation des résultats	A parfaire
5) Spécificité	Pathologies membre supérieur liées au travail
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	anglais
Coût	gratuit
7) Temps demandé	
Administration	5 minutes
Entraînement du testeur	non
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	anglais
Application dans une autre culture	à voir
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	+/-
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	Pas besoin
Intra observateur	Pas besoin
Test-retest	Non testée
Cohérence interne	Cronbach alpha 0.83 à 0.93.
3) Validité	
Contenu	Partielle
Construction	Partielle
Relation avec AVQ	Oui
Relation avec statut physique	non
Distinction des groupes	Oui
Critères	Non testée
Sensibilité au changement	Oui en partie
4) Normes	non

Tableau n°19: caractéristiques de l'UEFS.

V.4 Le questionnaire POS-Hand/Arm (Patient Outcomes of Surgery-Hand/Arm).

Ce questionnaire a été publié en 2004 par une équipe londonienne (18). Il a été conçu pour mesurer les résultats de la chirurgie de la main et du membre supérieur, avec une attention particulière pour l'avis du patient sur son traitement et pour le résultat esthétique. Macey en 1995 (67) avait déjà publié un questionnaire centré sur la main, intitulé « Patient Evaluation Mesure » (PEM), questionnaire qui prenait en compte l'avis du patient sur son traitement, sur l'hospitalisation et sur l'état général de sa main. Malheureusement ce questionnaire n'a fait l'objet d'aucune analyse statistique ni de publications, ce qui est dommage car il explorait des dimensions différentes de celles explorées par les précédents questionnaires.

- **Patients et méthodes :** pour la mise au point du questionnaire, les recommandations internationales ont été respectées. D'après les auteurs aucun questionnaire, même le DASH, n'avait fait l'objet d'une telle rigueur dans son développement.
  Stade 1 : choix des items. 75 questions choisies d'après une revue de la littérature, l'opinion d'experts et une interview auprès de 17 patients. Ces items ont été testés chez 23 patients.
  Stade 2 : réduction des items. Le questionnaire a été envoyé à 236 patients avant chirurgie et 262 patients après chirurgie, en moyenne 2 semaines avant la chirurgie et 3 mois après. Un groupe de 88 patients a reçu un deuxième questionnaire post chirurgie, 2 semaines après le premier pour évaluer la fiabilité test retest.
  Stade 3 : évaluation psychométrique. 196 questionnaires pré opératoires et 316 questionnaires post opératoires ont pu être analysés. Le POS a été comparé au DASH, au SF 36 et au Michigan hand questionnaire (MHQ).
- **Résultats :** les items ont été classés en 6 groupes : symptômes, activités physiques, fonctionnement psychologique, apparence cosmétique, complications et satisfaction. Les questions ont été posées en prenant en compte les 4 dernières semaines.
  La moitié des items ont été éliminés suite à l'analyse de réduction. Le questionnaire comprend 29 items en pré opératoire et 33 items en post opératoire en raison de l'adjonction d'une échelle de satisfaction. L'échelle de réponse va de 1 à 5. Le score total est ramené sur 100 par une règle de trois. Plus le score est élevé plus la santé est bonne, au contraire du DASH.
  Acceptabilité : elle est bonne avec peu de réponses manquantes et des réponses bien distribuées.
  Fiabilité : le Cronbach alpha est de 0.88 à 0.91 pour les 3 échelles en pré opératoire et de 0.76 à 0.94 pour les 4 échelles en post opératoire. La corrélation globale des items est de 0.38 à 0.87. La fiabilité test retest est bonne de 0.73 à 0.94 selon les sous échelles du questionnaire.
  Validité de construction : les inters corrélations entre les sous échelles du POS sont bonnes de 0.38 à 0.73. La comparaison avec les autres questionnaires montre des corrélations entre le POS et le DASH (r=0.63 à 0.88), le POS et le MHQ (r=0.19 à 0.77) et le SF 36 mais de façon très modérée.
  L'échelle activités physiques et symptômes du POS est hautement corrélée à l'échelle fonction/symptômes du DASH (r>0.80) et avec l'échelle travail et AVQ du MHQ (r>0.72).
  Sensibilité au changement : toutes les échelles du POS montrent des effets de tailles modérés à larges (0.47 à 1.08) et le POS différencie ceux qui s'améliorent de ceux qui ne s'améliorent pas, ce qui indique une bonne sensibilité au changement.

- **Discussion :** ce questionnaire a été bien étudié et est intéressant car il s'adresse à une population de patients opérés de la main et du bras, en s'intéressant de près à leur avis. Ses propriétés psychométriques sont bonnes, un peu en dessous de celles du DASH, avec lequel il est bien corrélé. Il existe d'ailleurs des domaines communs aux trois questionnaires (POS, DASH et MHQ), mais le POS est le seul qui évalue les patients en post chirurgical. Le POS pourrait être utilisé pour comparer des équipes chirurgicales, pour identifier des sous groupes de patients qui bénéficient plus de certaines techniques chirurgicales et enfin pour évaluer les bénéfices et les risques de différents traitements chirurgicaux.
  La limitation de l'étude est qu'elle n'a été faite que chez un collectif de patients anglais venant de trois hôpitaux et il faudrait que les résultats soient confirmés par d'autres études. La sensibilité au changement mérite d'être mieux étudiée.

En conclusion : le POS est un questionnaire bien construit avec des propriétés psychométriques qui sont bonnes mais qui demandent à être confirmées dans d'autres études. Il n'est disponible qu'en anglais pour le moment. Il est sûrement très intéressant dans une pratique chirurgicale. Le tableau ci-dessous résume ses propriétés.

PROPRIETES	POS
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Questionnaire d'auto évaluation
4) Interprétation des résultats	A parfaire
5) Spécificité	Population chirurgicale (pré et post op)
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	anglais
Coût	gratuit
7) Temps demandé	
Administration	?
Entraînement du testeur	non
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	anglais
Application dans une autre culture	à voir
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	+/-
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	Pas besoin
Intra observateur	Pas besoin
Test-retest	0.73 à 0.94 dans une étude
Cohérence interne	Cronbach alpha 0.76 à 0.94 dans une étude
3) Validité	
Contenu	Partielle
Construction	Oui mais dans une étude seulement
Relation avec AVQ	Oui
Relation avec statut physique	non
Distinction des groupes	Oui
Critères	Non testée
Sensibilité au changement	Oui en partie, mais dans une étude
4) Normes	non

Tableau n°20: caractéristiques du POS.

V.5 Le questionnaire PRWE (Patient Rated Wrist Evaluation).

V.5.1 Introduction.

Md McDermid, canadienne de l'Ontario a publié ce questionnaire en 1996 puis en 1998 (60,66). Il s'agit d'un questionnaire spécifique du poignet. Le but était de créer un questionnaire à la fois scientifique et pratique, utilisable comme mesure d'outcome du poignet. Il a fallu trois années pour arriver à la forme définitive du questionnaire. Nous verrons les différentes étapes de mise au point de ce questionnaire bien conçu, qui a l'immense avantage d'avoir été traduit en français.

V.5.2 Développement du PRWE.

Dans un premier temps, une étude par courrier auprès des 151 membres de l'IWI (International Wrist Investigators), groupe de thérapeutes spécialisés dans le poignet, a été pratiquée pour savoir ce que ces thérapeutes attendaient d'un nouvel outil, ce qu'il devait contenir comme items et sous quelle forme devait il être. 66 % des thérapeutes ont répondu. La douleur et la fonction dans les AVQ ont été les deux domaines retenus par les experts. Une liste d'items a été faite avec l'avis d'experts, de patients et d'après d'autres questionnaires. Cette première liste a été réduite après avis d'experts et test auprès de patients. Les items concernant la douleur ont été réduits pour couvrir tous les degrés de sévérité et pour s'intéresser aussi à la fréquence et à l'intensité. Les items sur la fonction devaient être effectués avec l'une ou l'autre des mains, par la majorité des patients et devaient être faciles à comprendre. Les items ont été limités à 5 pour la douleur et 10 pour la fonction. Une échelle de 0 à 10 pour la cotation a été choisie.

V.5.3 Propriétés psychométriques du PRWE.

- **Fiabilité** : la fiabilité test retest a été testée sur trois groupes de patients. Un groupe de 28 fractures de Pouteau-Colles aigues après la phase d'immobilisation, un groupe de 36 fractures de Pouteau-Colles traitées et un groupe de 35 fractures du scaphoïde à un an. Pour les groupes 1 et 2 le deuxième questionnaire a été envoyé 2 à 7 jours plus tard. Pour le groupe 3 le deuxième questionnaire a été envoyé à un an d'évolution et administré en même temps que le SF36.
  Une excellente fiabilité pour les trois groupes est retrouvée avec un ICC> 0.90. La seule sous échelle du SF 36 qui a une excellente corrélation avec le PRWE est la fonction physique avec un ICC=0.89.
- **Validité :**
  La validité de construction était basée sur le fait que les patients ayant une fracture du poignet, doivent avoir moins de douleur et d'incapacité au cours du temps. 101 patients ayant une fracture du poignet ont rempli le questionnaire à T0 puis 2, 3 et 6 mois après la fracture. Une analyse ANOVA (ANalysis Of Variance) a été faite. Cette technique statistique permet de comparer les moyennes de plus de deux populations.Une amélioration statistiquement significative du score PRWE est apparue au cours du temps (p<0.0001). Le score moyen est passé de 74 à T0 à 19 à 6 mois. Le sous test physique du SF 36 a suivi le même chemin
  La validité de critères a été calculée en comparaison avec le SF36 et un score composite formé des mobilités articulaires du poignet, de la force et du test de Jebsen

(30 points mobilités, 40 points force et 15 points dextérité). Les corrélations étaient élevées entre le score PRWE douleur et le SF36 douleur (r=0.72). Les corrélations entre le score PRWE et le SF36 physique étaient plus élevées qu'avec le SF 36 mental (0.63 versus 0.33). Les corrélations avec le score composite étaient modérées (r=0.52). Toutes ces corrélations montrent l'association entre les différentes mesures et renforce le fait que le PRWE mesure bien ce pour quoi il a été conçu.

- **Sensibilité au changement :** deux études de McDermid, dont nous avons déjà parlé lors du chapitre sur le DASH ont étudié ce paramètre (62, 65). La première (62) concernait 59 patients avec fracture du radius distal évalués à T0 puis à 3 et 6 mois. Le PRWE a été comparé au DASH et au SF36. Le PRWE était le plus sensible des trois avec une SRM de 2,27. Il était surtout sensible dans les trois premiers mois de la fracture lorsque les tests physiques ne pouvaient être faits. Le même score composite que celui décrit au paragraphe précédent a montré qu'il était plus sensible que le questionnaire dans la période 3 à 6 mois, surtout en ce qui concerne la force du grip. Cette étude montre que le questionnaire le plus spécifique du poignet est le plus sensible pour évaluer l'évolution de la fracture du poignet. La deuxième étude (65) concernait 60 patients (36 mains et 24 poignets) avec un questionnaire PRWE modifié, auquel ont été ajouté deux questions sur l'apparence de la main. Le questionnaire PRWE était comparé au DASH à T0 et à 3 mois. Comme dans la première étude, le PRWE a été légèrement plus sensible que le DASH avec une SRM de 1.51 contre 1.37 pour le DASH. Les pathologies représentées pour le poignet étaient des fractures du poignet, des instabilités du carpe et de l'arthrose. Pour la main il s'agissait de fractures, de lésions tendineuses, de Dupuytren opérés et de prothèses des inter phalangiennes. Il existait une forte relation entre le changement du score DASH et du PRWE (r=0.72) mais pas de corrélation entre la question supplémentaire et les deux questionnaires.

V.5.4 Autres études.

- McDermid a publié deux autres études sur le PRWE concernant le suivi à un an de fractures du poignet (63,64). La première (63) a suivi 275 patients pendant un an de façon prospective en utilisant le PRWE, le DASH et le SF 36. Les buts étaient multiples. Tout d'abord, avoir des résultats descriptifs sur une grande série, de différentes données cliniques et des questionnaires sus nommés. Ensuite de décrire l'histoire naturelle de la récupération de la fracture du poignet. Enfin, les auteurs ont essayé de savoir si avec les résultats à T0 il pouvait prédire l'évolution à un an. Seuls les résultats de 250 patients sont interprétables. Pour les questionnaires, les résultats sont similaires aux études citées plus haut (62,65) en terme de sensibilité au changement Il a été montré dans cette étude que la compensation financière à un an entraîne une augmentation du score PRWE de 18 points, par rapport aux patients non compensés. Cet effet n'a pas été retrouvé pour le DASH et le SF36. Le facteur prédictif important semble le score élevé à T0 et ceci pour les 3 questionnaires, mais seul 21 à 27 % du score final peut être expliqué par ces modèles prédictifs. La deuxième étude (64) a suivi aussi pendant un an 129 patients ayant une fracture du radius. Le but de l'étude était d'étudier la douleur et l'incapacité avec le questionnaire PRWE chez ces patients et de voir comment ces deux variables changeaient avec le temps. La majorité des patients s'amélioraient dans les six premiers mois, et une minorité a des difficultés importantes à un an. Cette base de données peut permettre en comparant les scores moyens avec le score du patient, de détecter une éventuelle complication telle une algodystrophie débutante.

- Karnezis (50) rapporte une étude prospective chez 31 patients avec fracture du radius instable, dont le but est de rechercher une association entre le PRWE et des variables cliniques objectives. Les données sont recueillies à 6 semaines (après ablation du plâtre), puis à 3, 6, 12 et 24 mois. Les variables étudiées sont le PRWE, les mobilités du poignet (flexion/extension, prono-supination) et la force au Jamar. L'analyse statistique a été faite en régression multiple en formulant un score de risque avec 4 variables (age, sexe, traumatisme à haute énergie et une fracture intra articulaire). Les variables cliniques mesurées étaient considérées comme des possibles variables pronostiques. Le résultat est que quand la force au Jamar augmente le score PRWE diminue. Mais d'autres facteurs non mis en évidence dans cette étude influencent le score. L'étude n'est pas bien étayée sur le plan statistique. Les auteurs concluent que toutes les variables cliniques du poignet ne reflètent pas le niveau global de la fonction du poignet de la même manière. Dans cette étude les mobilités n'étaient pas associées à l'amélioration du score PRWE.
- Adams (2) a mené une étude originale avec 21 sujets normaux qu'il a équipé d'orthèses de restriction du poignet. Le but était de quantifier et de comparer l'incapacité causée par cette restriction de mouvement soit totale, soit partielle (30° de flexion extension). Le DASH et le PRWE ont été administrés ainsi que le test de Jebsen et une série de 13 AVQ sélectionnées. Les scores DASH et PRWE étaient significativement plus élevés quand le poignet était partiellement ou complètement limité ($p<0.05$). Les différences entre le blocage partiel ou total étaient aussi significatives. Les auteurs concluent que l'incapacité perçue (par les questionnaires) lors d'une réduction de la mobilité du poignet apparaît plus grande que lorsqu'elle est mesurée par l'examen clinique classique. Les deux questionnaires ont montré encore une bonne sensibilité au changement.

V.5.5 Version française du PRWE et instructions pour sa cotation.

Cette version a été publiée en 2003 par l'équipe du Dr Dubert qui avait déjà fait la traduction du DASH (93). Le PRWE n'étant pas une marque déposée, il n'y a pas besoin de faire une traduction et des contre traductions comme pour le DASH. De plus, les questions sont assez proches de celles du DASH. Quelques adaptations ont été faites mais cette version française respecte la présentation générale et la mise en forme du questionnaire original. Le questionnaire français a été testé chez 20 patients dont 15 étaient traités pour fracture du radius. Le temps moyen de remplissage a été de 4 minutes. Comme le DASH, le score le plus bas traduit le meilleur résultat et inversement. Le questionnaire et la façon de l'interpréter sont tirés de l'article de Dubert et collaborateurs.

- **Questionnaire d'évaluation de votre poignet atteint:PRWE.**

Nom et Prénom :
……………………………………………………………………………………………….…
………..

Age :
…………………………………………………………………………………………………
…………………………..….

Etes vous : Droitier (ère) ☐ Gaucher(ère) ☐

Votre poignet atteint : Droit ☐ Gauche ☐

Votre profession :
………..

Date :
……..

*Les questions suivantes nous aideront à évaluer le niveau de difficulté que vous avez eu avec votre poignet **au cours de la semaine passée**. Pour chaque question vous devez décrire la gêne moyenne rencontrée avec votre poignet durant la semaine passée sur une échelle de 0 à 10.*
SVP, essayer de répondre à **TOUTES LES QUESTIONS**.
Si vous n'avez pas effectué une des activités demandées, veuillez **ESTIMER** la douleur ou la difficulté que vous auriez eu à réaliser cette activité.
Si vous n'avez jamais effectué une des activités demandées, vous pouvez ne pas répondre à la question et laissez la ligne vierge.

1. **DOULEUR**.

*Pour chaque question, marquez le niveau moyen de douleur ressentie au niveau de votre poignet **au cours de la semaine passée** en **ENTOURANT le chiffre (UN SEUL CHIFFRE)** qui correspond le mieux à votre douleur sur une échelle de 0 à 10. Le zéro correspond à aucune douleur et le 10 à la plus forte douleur que vous ayez ressentie en réalisant l'activité demandée.*

Echelle exemple (ne rien entourer) 0 1 2 3 4 5 6 7 8 9 10
Aucune douleur Douleur extrême

EVALUEZ VOTRE DOULEUR

Au repos	0	1	2	3	4	5	6	7	8	9	10
En effectuant une tâche nécessitant des mouvements répétés du poignet	0	1	2	3	4	5	6	7	8	9	10
En soulevant un objet lourd	0	1	2	3	4	5	6	7	8	9	10
Lorsque la douleur est à son maximum	0	1	2	3	4	5	6	7	8	9	10
Avez vous souvent mal ? (Fréquence de votre douleur)	0 Jamais	1	2	3	4	5	6	7	8	9	10 tout le temps

2. EVALUEZ VOTRE FONCTION

a. Activités spécifiques.

Pour chaque question, marquez le niveau de difficulté rencontrée ***<u>au cours de la semaine passée</u>*** *pour réaliser l'activité demandée en* ***ENTOURANT*** *le chiffre* ***(UN SEUL CHIFFRE)*** *qui correspond le mieux à votre difficulté sur une échelle de 0 à 10. Le zéro correspond à l'absence totale de difficulté et le 10 à l'incapacité complète à réaliser l'activité demandée.*

Echelle exemple (ne rien entourer)	0 Aucune Difficulté	1	2	3	4	5	6	7	8	9	10 Incapacité à réaliser

Tourner la poignée d'une porte avec votre main atteinte.	0	1	2	3	4	5	6	7	8	9	10
Couper la viande en utilisant votre main atteinte.	0	1	2	3	4	5	6	7	8	9	10
Fermer les boutons de votre chemise.	0	1	2	3	4	5	6	7	8	9	10
Utiliser votre main atteinte pour vous lever d'une chaise ou d'un fauteuil.	0	1	2	3	4	5	6	7	8	9	10
Porter un objet de 5 kg avec votre main atteinte	0	1	2	3	4	5	6	7	8	9	10
Utiliser une serviette de toilette avec votre main atteinte.	0	1	2	3	4	5	6	7	8	9	10

b. Activités courantes.

Pour chaque question, marquez le niveau de difficulté rencontrée ***<u>au cours de la semaine passée</u>*** *pour réaliser l'activité demandée en* ***ENTOURANT*** *le chiffre* ***(UN SEUL CHIFFRE)*** *qui correspond le mieux à votre difficulté sur une échelle de 0 à 10. Par activités courantes nous entendons les activités telles que vous les réalisiez avant d'avoir vos problèmes au poignet. Le zéro correspond à l'absence totale de difficulté et le 10 à l'incapacité complète à réaliser l'activité demandée.*

Echelle exemple (ne rien entourer)	0 Aucune difficulté	1	2	3	4	5	6	7	8	9	10 Incapacité à réaliser

Hygiène et soins personnels (S'habiller,se laver)	0	1	2	3	4	5	6	7	8	9	10

Effectuer vos travaux ménagers à domicile (ménage, entretien).	0 1 2 3 4 5 6 7 8 9 10
Effectuer votre travail (activité professionnelle ou travail quotidien habituels)	0 1 2 3 4 5 6 7 8 9 10
Effectuer vos activités de loisirs	0 1 2 3 4 5 6 7 8 9 10

- **Mode d'emploi du questionnaire sur le poignet PRWE.**

Le PRWE comporte 15 questions courtes : 5 questions sur la douleur et 10 questions sur la fonction. 4 questions sur la douleur portent sur l'intensité et 1 question sur la fréquence.
6 questions sur la fonction portent sur des gestes spécifiques dont 3 réalisés uniquement avec le poignet et la main atteints et 4 questions sur des activités plus générales de la vie quotidienne.
La réponse se fait selon le principe d'une échelle numérique de 1 à 10.
Le patient doit remplir seul le questionnaire avec une seule réponse par question. L'évaluation porte sur les 7 derniers jours. Si le patient n'a pas effectué l'activité au cours de la semaine passée, il doit estimer sa capacité. La réponse est laissée vierge seulement si le patient n'a jamais réalisé l'activité demandée.

Calcul du score :
Trois modalités de calcul du score total sont possibles.

1) *Analyse globale* : l'addition des chiffres obtenus rapporte un score sur 150.
2) *Analyse combinée* : le score de la fonction est divisé par 2 puis additionné à celui de la douleur afin d'obtenir un score sur 100.
3) *Analyse séparée de la douleur et de la fonction* : le score sur la douleur est multiplié par deux afin d'obtenir un score sur 100, celui de la fonction est interprété tel quel.

Ce questionnaire se présente comme une combinaison de 3 groupes d'items pouvant être évalués séparément si besoin. La validité et la fiabilité des sous-groupes ont été analysées et confirmées. Néanmoins, la fiabilité du score global ou combiné est supérieure à celle des scores séparés.
En cas d'absence de réponse, le Dr Dubert suggère de calculer et de rapporter le score au nombre de questions remplies selon une simple règle de trois.
Exemple : un score global de 52 obtenu avec 13 questions remplies (maximum 130) correspond à un score de $60 = \frac{15 \times 52}{130}$.
Sur un total de 15 questions, on peut accepter au maximum 2 questions vierges pour garder au questionnaire sa validité.

V.5.6 Conclusions.

Le questionnaire PRWE est un questionnaire facile à employer, avec peu d'items. Il peut être rempli en 5 minutes dans une salle d'attente par exemple. Il a de bonnes propriétés psychométriques en ce qui concerne la fiabilité et la validité. La cohérence interne n'est pas

connue. La sensibilité au changement semble bonne mais demande à être confirmée dans d'autres études faites par d'autres équipes.
Il a été bien étudié dans la fracture du radius. Dans d'autres pathologies, son intérêt demande à être confirmé. Le grand avantage pour les francophones c'est qu'il existe une version française utilisable facilement. Il peut être utilisé, en même temps ou à la place du DASH dans une pathologie de la main ou du poignet. Je pense qu'il a toute sa place dans une pratique clinique de tous les jours au même titre que l'examen clinique dont il est le complément.
Le tableau ci-dessous résume ses caractéristiques.

PROPRIETES	PRWE
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Questionnaire d'auto évaluation
4) Interprétation des résultats	A parfaire
5) Spécificité	Pathologies poignet +/- main
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	anglais
Coût	gratuit
7) Temps demandé	
Administration	5 minutes
Entraînement du testeur	non
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	Anglais et français
Application dans une autre culture	oui
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	+/-
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	Pas besoin
Intra observateur	Pas besoin
Test-retest	> 0.90
Cohérence interne	?
3) Validité	
Contenu	Partielle
Construction	Oui
Relation avec AVQ	Oui
Relation avec statut physique	Non
Distinction des groupes	Non
Critères	Oui
Sensibilité au changement	Oui dans deux études
4) Normes	non

Tableau n° 21: caractéristiques du PRWE.

V.6 Le Michigan Hand outcomes Questionnaire (MHQ).

V.6.1 Introduction.

Chung, chirurgien plasticien a publié ce questionnaire en 1998 et 1999 dans deux articles (21, 22). Le but des auteurs était de développer un questionnaire qui pourrait être utilisé de manière standardisée pour mesurer l'outcome chez des patients avec tous types de pathologies de la main, surtout en chirurgie de la main.

V.6.2 Développement du MHQ.

Le MHQ préliminaire s'est inspiré du SF 36 et de l'AIMS. Le McGill pain questionnaire et le questionnaire sur le canal carpien (Boston questionnaire) ont été utilisés pour le développement d'une liste de questions pour quantifier la douleur. Des patients ont été consultés et ont ajoutés des items qu'ils jugeaient importants. La première version comportait 100 items. Un panel comprenant des patients, des thérapeutes et des chirurgiens de la main ont classés ces questions en 6 échelles : fonction globale de la main, AVQ, douleur, performance au travail, esthétique et satisfaction du patient sur la fonction de la main. Ensuite en utilisant la technique statistique d'analyse de facteur, les items ont été réduits à 37 dans 6 domaines différents. Sur les 6 échelles, 4 contiennent des items mesurant la main droite et la main gauche, ceci afin de contrôler l'effet confondant de la main dominante. Pour le calcul du score, on additionne les réponses, qui vont de 1 à 5 (comme le DASH) et on ramène le tout sur 100. Comme le DASH, plus le score est élevé plus la fonction est mauvaise.

V.6.3 Propriétés psychométriques.

Le questionnaire a été administré avec le SF12 à 200 patients consécutifs avec des symptômes depuis au moins trois mois. Pour la fiabilité test retest, 25 patients ont rempli le questionnaire 7 jours après.

- **Fiabilité :** la fiabilité test retest a été excellente avec dans 5 des 6 échelles une corrélation supérieure à 0.85. La différence moyenne des scores entre deux administrations a été calculée. Cette différence était inférieure à 5 points avec un intervalle de confiance à 95 % proche de 0 ce qui va en faveur d'une très bonne fiabilité test retest.
  La cohérence interne évaluée par le Cronbach alpha va de 0.86 à 0.97 dans chacune des 6 échelles.
- **Validité :** la validité de construction était basée sur l'hypothèse que les 6 échelles du MHQ étaient corrélées entre elles. Les corrélations de Spearman ont été calculées. Elles sont élevées (entre 0.63 et 0.78) pour les 5 échelles qui mesurent la partie fonctionnelle : fonction globale, AVQ, travail, douleur et satisfaction avec la fonction de la main. L'échelle fonction globale de la main est hautement corrélée avec les autres échelles. Les patients qui ont une douleur importante ont des mauvais résultats dans les autres échelles. L'échelle esthétique a des corrélations très modérées avec les autres échelles. Pour déterminer quelle échelle du MHQ est un prédicteur de la fonction physique, une régression des échelles du MHQ contre le composant physique du SF12 a été pratiquée. L'échelle douleur est prédicteur de la fonction physique ($p<0.001$). Les autres échelles ne sont pas significatives. La validité de contenu est assurée par le développement du questionnaire et l'avis des experts. La validité de critères n'a pas été testée.

- **Sensibilité au changement :** la même équipe a conduit cette étude sur 187 patients avec des mains chroniques. Le MHQ a été administré à T 0 puis par courrier 6 à 18 mois après. 92 patients ont renvoyé le questionnaire soit 49 %. Le deuxième questionnaire était identique au premier sauf qu'il y avait une question additionnelle après chaque échelle, destinée au patient qui devait apprécier son changement par rapport au premier questionnaire. Les corrélations de Spearman ont été utilisées pour corréler les réponses du patient à la question supplémentaire et le changement du score entre les deux questionnaires. Les deux groupes répondeurs et non répondeurs étaient identiques. Les corrélations entre l'avis du patient et le changement du score pour les 6 échelles sont modérées (entre 0.25 pour l'esthétique et 0.43 pour la douleur) avec un $p< 0.05$. Il faut noter que 80 % des patients se disent mieux et 5 à 10% se disent pire. Ce résultat un peu surprenant peut être lié soit à un biais d'évocation (délai d'au moins 6 mois entre les 2 questionnaires) soit est relatif à la perception individuelle sur l'amplitude du changement. Certains patients avec de grandes attentes peuvent percevoir une grande amélioration dans le score MHQ comme un petit effet. L'explication la plus probable est que la méthode statistique utilisée ici, qui est la méthode heuristique, n'est pas la meilleure et qu'il faudrait utiliser les SRM et l'effet de taille. Il aurait été intéressant aussi de comparer le MHQ au DASH. D'autres études sont donc indispensables pour confirmer cette sensibilité au changement.
- Une autre étude a comparé le MHQ au POS, c'est celle de Cano (18) dont j'ai déjà parlé auparavant. Les activités physiques et les symptômes du POS étaient hautement corrélés au MHQ AVQ et travail ($r>0.72$). La corrélation était bonne entre le POS symptômes et le MHQ douleur ($r>0.77$). Pour les autres échelles les corrélations étaient plus faibles aux environ de 0.30.
- En 2005, une étude de Kotsis (52), a comparé la sensibilité au changement du MHQ et du DASH après chirurgie du canal carpien. Il s'agit d'une étude prospective de 50 patients qui ont remplis les 2 questionnaires avant et 6 mois après chirurgie. Tous les domaines du MHQ s'améliorent significativement après chirurgie. La sous échelle douleur du MHQ a la plus grande SRM à 0.9 et la sous échelle fonction a une SRM de 0.6. La SRM du DASH est de 0.7. Le MHQ a donc une bonne sensibilité au changement, comparable au DASH après chirurgie du canal carpien.

V.6.4 Conclusions.

Le MHQ est un questionnaire intéressant qui explore d'autres dimensions que le DASH par exemple. Il est assez proche du POS. Ses propriétés psychométriques demandent à être confirmées par des études plus nombreuses et surtout faites par d'autres équipes. La sensibilité au changement semble discutable en l'état actuel des données. Le calcul du score parait un peu compliqué. Je pense qu'il est un peu prématuré de l'utiliser de façon large tant que ses propriétés n'auront pas été mieux précisées. Il n'est disponible qu'en anglais. Le tableau ci-dessous résume ses caractéristiques.

PROPRIETES	MHQ
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Questionnaire d'auto évaluation
4) Interprétation des résultats	A parfaire
5) Spécificité	Pathologies main / poignet
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	anglais
Coût	gratuit
7) Temps demandé	
Administration	10 minutes
Entraînement du testeur	non
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	Anglais
Application dans une autre culture	Oui a priori
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	+/-
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	Pas besoin
Intra observateur	Pas besoin
Test-retest	>0.85 dans une étude
Cohérence interne	0.86 à 0.97 dans une étude
3) Validité	
Contenu	Oui
Construction	Oui en partie
Relation avec AVQ	Oui
Relation avec statut physique	Non
Distinction des groupes	Non
Critères	Pas étudiée
Sensibilité au changement	Oui dans une étude corrélations modérées
4) Normes	non

Tableau n°22: caractéristiques du MHQ.

V.7 ABILHAND.

ABILHAND a été publié en 1998 par une équipe belge et italienne (74). C'est une mesure de la capacité manuelle. Les auteurs soulignent en préambule que la relation entre l'altération (dysfonction au niveau de l'organe) et l'incapacité (dysfonction au niveau de la fonction) n'est pas franche. D'après les auteurs, les tests existants sont trop artificiels et ne reposent pas assez sur des activités de la vie quotidienne. Le temps comme mesure de la capacité dans la plupart de ces tests, n'est pas discriminant. Le but des auteurs a donc été de construire une échelle validée de la capacité manuelle. Des patients ayant une arthrodèse du poignet dans le cadre d'une PR ont été pris comme population de référence en raison de leur homogénéité. Le modèle de Rasch, dont nous avons déjà parlé, a été utilisé pour valider ce score. Il s'agit d'un questionnaire auto administré et auto évalué.

- **Patients et méthodes :** 18 patients ayant une arthrodèse du poignet dans le cadre d'une PR. Ils étaient tous droitiers et ont été évalués 6.8 à 40.9 ans après le diagnostic de PR et 0.5 à 16.8 ans après l'arthrodèse. Le questionnaire ABILHAND est constitué de 57 items avec des activités uni et bi manuelles. Les items sont tirés en partie d'échelles existantes. Les items sélectionnés le sont pour que la dextérité manuelle soit vitale pour la réussite de l'action.
  On demande aux patients d'estimer leur facilité ou leur difficulté à faire chaque activité sur une échelle de 0 à 5 : 0= impossible, 1=très difficile, 2=difficile, 3=assez facile, 4=facile, 5=très facile. Si l'activité n'a jamais été faite, le patient ne répond pas. Chaque item est lu par l'examinateur au patient. Pour les activités uni ou bi manuelles, on demande au patient de répondre sans prendre en compte le bras utilisé pour faire l'activité. Pendant l'évaluation les auteurs se sont rendus compte de la difficulté des patients à se placer dans les 6 catégories. 4 catégories ont été retenues alors, 0=très difficile, 1=légèrement difficile, 2=facile, 3=très facile.
  L'analyse a été faite selon le modèle de Rasch, qui est un modèle ordonnateur plutôt que descriptif. Il s'agit d'une analyse très complexe difficile à résumer. Le but est d'arriver à avoir un outil unidimensionnel (cf chapitre II).
- **Résultats :** 46 des 57 items définissent une capacité manuelle commune unique. Les items en rapport avec l'alimentation, l'hygiène et l'habillage du haut du corps sont bien corrélés aux items d'autres échelles qui explorent les mêmes activités. La fiabilité de l'échelle est de 0.90, dans le modèle de Rasch, il n'y a pas besoin de répéter le test pour la calculer. ABILHAND est donc unidimensionnel.

Une autre étude (citée dans 8) a étudiée ABILHAND chez des patients avec des TCC (traumatismes cranio cérébraux) chroniques. Il a été modifié en rajoutant des activités bi manuelles qui semblaient plus discriminatives dans cette population de TCC. A noter que dans cette étude l'échelle de réponse des patients a été encore modifiée pour n'avoir plus que trois niveaux : 0=impossible, 1= difficile, 2=facile.

Je n'ai pas retrouvé d'autres publications concernant ce questionnaire. Ce questionnaire serait traduit en français, anglais, italien et suédois. Pour ma part je pense qu'il s'agit d'un concept compliqué, avec une méthode statistique inhabituelle sans comparaison avec une autre méthode (par exemple analyse en ACP). Je ne vois pas ce qu'il peut apporter de plus, en l'état actuel des connaissances, que les questionnaires déjà existants et mieux étudiés.

V.8 Le questionnaire de Boston (BQ) sur le canal carpien.

V.8.1 Introduction.

Ce questionnaire a été publié en 1993 par Levine et collaborateurs (56) à Boston à l'hôpital de Brigham. Il est connu soit sous le nom de Boston questionnaire (BQ) soit sous le nom de questionnaire de Brigham, les deux terminologies étant retrouvées dans la littérature.
Les résultats de l'intervention sur le syndrome du canal carpien (SCC) sont généralement évalués avec des données physiques, mais pas sur les données concernant les symptômes et la fonction qui intéressent plus les patients. Les auteurs ont donc développé un questionnaire auto administré pour mesurer la sévérité des symptômes et le statut fonctionnel des patients ayant un SCC. Ce questionnaire nous le verrons est devenu la référence pour cette pathologie.

V.8.2 Développement du questionnaire.

Un panel de chirurgiens de la main, de rhumatologues et de patients a identifié 6 domaines pour l'évaluation du SCC : douleur, paresthésies, engourdissements, faiblesse, symptômes nocturnes et le statut fonctionnel global. Une échelle de sévérité des symptômes comprenant ces 6 items a été crée. Elle comprend 11 questions à choix multiples, cotées de 1 à 5. Le score de sévérité des symptômes est calculé comme la moyenne des scores aux 11 questions.
12 activités fonctionnelles souvent touchées par le SCC ont été identifiées. 4 ont été éliminées et les 8 activités restantes ont constitué l'échelle du statut fonctionnel, cotées de 1 (pas de difficulté) à 5 (impossible à faire). Le score de l'échelle du statut fonctionnel est la moyenne des 8 scores des items. Les 8 activités choisies sont :

1. Ecrire.
2. Boutonner des habits
3. Tenir un livre en lisant.
4. Prendre un combiné de téléphone.
5. Ouvrir un pot.
6. Faire le ménage.
7. Porter des sacs à provision.
8. Se laver et s'habiller.

Deux groupes de patients ont ensuite été suivis. Une cohorte prospective de 67 patients avec SCC. 39 ont été évalués avant une intervention et 28 ont été traités médicalement. Les premiers 31 patients ont servis à mesurer la fiabilité test restest. La cohérence interne a été calculée sur les 67 patients. 43 patients ont permis d'analyser la validité.
Les 39 patients qui devaient être opérés, ont formé une cohorte prospective d'outcome après chirurgie du canal carpien. A la date de publication de l'étude, seuls 26 patients peuvent être analysés.

V.8.3 Propriétés psychométriques.

- **Fiabilité test retest :** elle est excellente avec un coefficient de corrélation de Pearson de 0.91 pour l'échelle des symptômes et de 0.93 pour l'échelle de la fonction.
- **Cohérence interne :** le Cronbach alpha est de 0.89 pour l'échelle des symptômes et de 0.91 pour l'échelle de la fonction, ce qui traduit une corrélation inter item élevée dans chaque échelle.
- **Validité :** comme toujours, elle est difficile à apprécier. Les scores ont été comparés à des variables cliniques : force au grip et de la pince, conduction sensitive du nerf

médian et résultats du test de Semmes-Weinstein et de discrimination de deux points. L'hypothèse de départ était que les échelles devaient avoir des corrélations positives mais modérées avec ces mesures traditionnelles, car ces échelles mesurent l'incapacité qui n'est en général pas mesurée par l'examen physique.
Le score de sévérité des symptômes est de 3.0 +- 0.96 points (moyenne et écart type) et celui du statut fonctionnel de 2.65 +- 0.96 points. Toutes les corrélations vont dans le même sens, à savoir les plus mauvais scores sur les 2 échelles sont associés à une altération plus sévère. Le score de statut fonctionnel a une haute corrélation avec la sévérité des symptômes (r=0.63 avec $p < 0.001$) indiquant que les patients ayant des symptômes sévères ont des limitations fonctionnelles majeures. Il existe une corrélation modérée entre la force au grip et à la pince et le statut fonctionnel (r= 0.50 et 0.60 avec $p<0.001$) et une corrélation faible avec la sévérité des symptômes (r=0.38 et 0.47 avec $p< 0.05$ et 0.01).

- **Sensibilité au changement :** deux cohortes de patients ont été utilisés pour cela. Une première cohorte de 38 patients opérés en 1990, évalués 14 mois après de manière rétrospective. Ils remplissaient un questionnaire sur leur état actuel et devaient se rappeler de leur état pré opératoire pour remplir l'autre questionnaire. La sensibilité a été appréciée par l'effet de taille. Les patients devaient aussi noter sur une échelle de 5 points leur satisfaction par rapport à l'intervention. Les changements de scores ont été corrélés à la satisfaction en utilisant le coefficient de corrélation de Spearman. La deuxième cohorte de 26 patients a rempli le questionnaire avant et 3 mois après l'intervention.
  Le tableau ci-dessous résume les résultats des 2 cohortes **(Tableau n°23)**

	Scores pré opératoires	Scores post opératoires	Effet de taille	Corrélations avec la satisfaction du patient
COHORTE 1 38 patients	3.4 + - 0.67 points (symptômes) 3.0 +-0.93 points (fonction)	1.9 +- 1 points (symptômes) 20 +- 1,1 points (fonction)	1.4 0.82	r= 0.52 (p=0.0007) r= 0.29 (p=0.09)
COHORTE 2 26 patients	3.1 +- 0,9 points (symptômes) 2.7 +-1.0 points (fonction)	2.0 +- 1,0 points (symptômes) 2.1 +-1,1 points (fonction)	1.13 0.71	r=0.50 (p<0.01) r=0.54 (p<0.01)

Ces résultats vont tous dans le même sens, à savoir une sensibilité au changement du questionnaire. L'amélioration est substantielle et survient dans les trois premiers mois ce qui est concordant avec l'expérience clinique. Ce questionnaire est plus reproductible que la plupart des mesures cliniques utilisées habituellement pour le suivi du SCC. Ces résultats ont donc confirmé l'hypothèse de départ émise pour la validité du questionnaire.

Les limitations de cette étude sont :

- Une population hétérogène de SCC (jeunes, âgés, travailleurs manuels ou non) ce qui peut entraîner un manque d'homogénéité dans les résultats.
- Qu'il n' y a que 13 % des patients qui reçoivent une compensation financière et qu'il n'a donc pas été possible de comparer si l'évolution était différente entre ceux qui sont compensés et ceux qui ne le sont pas.
- Caractère unicentrique de l'étude.
- Les résultats de la première cohorte sont rétrospectifs.

V.8.4 Autres études.

Depuis la publication initiale, de nombreuses autres publications se sont intéressées au BQ. Je n'ai sélectionné que celles qui me semblaient intéressantes et qui comparaient le BQ à d'autres questionnaires. Le BQ a été traduit en italien en 1998 (cité dans 73), en suédois, en portugais et en espagnol. Je n'ai pas retrouvé de trace d'une version française.

- Rosales (76) en 2003 a publié la version espagnole du BQ en même temps que celle du DASH. Il a étudié 42 patients ayant un SCC et leur a administré le DASH et le BQ. La fiabilité test retest du BQ espagnol est bonne avec un coefficient de corrélation de Pearson de 0.87 et 0.85 respectivement pour les 2 échelles du BQ. La cohérence interne montre un Cronbach alpha à 0.90 et 0.91. Ceci confirme les résultats de la version initiale et la bonne qualité de la traduction.
- Gay (34) toujours en 2003, a comparé la sensibilité au changement du BQ, du DASH et du SF36 après chirurgie du SCC chez 34 patients. Les questionnaires étaient administrés avant puis à 6 et 12 semaines post opératoires. Comme on pouvait s'y attendre, la sensibilité au changement du BQ a été supérieure à celle du DASH qui est moins spécifique du SCC. L'effet de taille et les SRM étaient respectivement de 1.71 et 1.66 pour le BQ et de 1.01 et 1.13 pour le DASH. La corrélation entre le changement des scores des deux questionnaires était très bonne avec un coefficient de corrélation de Spearman à 0.87 à 12 semaines. La validité concurrente est donc bonne avec le DASH. Le seul bémol dans cette étude est qu'il n'a pas été retrouvé de corrélation entre le changement des scores du BQ et du DASH avec la satisfaction du patient alors même que 85 % des patients se disaient très ou complètement satisfaits à 6 semaines. Ce manque de corrélation ne prouve pas que la variation du score du BQ n'a pas de signification clinique, mais il faut plutôt attribuer cela au mode de calcul statistique utilisé qui ne serait pas la méthode la plus appropriée.
- Greenslade (38) en 2004 a aussi étudié la sensibilité au changement du BQ et du DASH chez 57 patients ayant un SCC, en pré opératoire et à 3 mois post opératoire. L'étude était conçue pour étudier surtout le DASH. L'échelle fonction du BQ avait une sensibilité comparable au DASH (SRM de 0.62 versus 0.66), mais l'échelle symptômes avait une sensibilité supérieure au DASH (SRM de 1.07 versus 0.66). Il existait une corrélation entre l'importance du score pré opératoire et l'amélioration post opératoire. La fiabilité test retest du BQ a été calculée chez 31 patients; elle était de 0.82 pour le BQ symptômes et 0.79 pour le BQ fonction ce qui est comparable aux autres études. Le temps moyen de remplissage du BQ a été de 5,6 minutes +- 3,5 minutes.
- Amadio (6), en 1996, a utilisé le BQ dans l'outcome après fracture du poignet, en le comparant au SF36, à l'AIMS et à des variables cliniques (force, mobilités, sensibilité à la pression et dextérité). 21 patients ont été inclus et les mesures ont été faites à T0 et à 3 mois. La seule modification dans le BQ a été de remplacer le mot canal carpien par fracture. Des deux échelles du BQ, seule l'échelle fonction a montré une amélioration significative à 3 mois. Les SRM sont respectivement de 0.36 pour le BQ symptômes et 0.84 pour le BQ fonction. Dans cette étude aucunes corrélations n'ont été trouvées entre les questionnaires et les variables cliniques, sauf entre une sous échelle de l'AIMS et la force de la pince à 3 mois. Les auteurs n'apportent pas d'explications claires à ces résultats, en dehors du fait que l'échantillon est petit et qu'il n'y a pas de groupe contrôle. On peut tout de même remarquer que le choix du BQ comme outil d'outcome pour une fracture du poignet n'est peut être pas très judicieux étant donné sa spécificité. D'ailleurs seule l'échelle fonction du BQ s'est améliorée. L'échelle

symptômes étant très spécifique du SCC, on peut comprendre son absence de réponse au cours d'une fracture du poignet.

- Jain (44), en 2001 a étudié le BQ, le DASH et le SF36 chez 43 patients opérés d'un conflit ulno carpien. Les questionnaires ont été envoyés par courrier. Il s'agit d'une étude rétrospective. Le BQ et le DASH ont une distribution similaire des réponses. Le BQ et le DASH ont une bonne corrélation entre eux (r=0.85) et corrèlent de façon modérée avec le SF36 fonction physique et douleur (r entre 0.46 et 0.68). les corrélations avec le SF36 composant mental sont faibles (r=0.38 et 0.39). le BQ et le DASH ont été capables de discriminer trois groupes de patients : pas de problèmes ou problèmes moyens, problèmes modérés ou sévères et problèmes très sévères. (p=0.0001 pour le BQ). Les patients qui rapportent pas ou des problèmes moyens ont des scores plus faibles que ceux qui rapportent des problèmes modérés ou sévères. Le BQ a été capable de discriminer les patients qui travaillent avec restrictions dues à leur poignet de ceux qui travaillent sans restrictions (p=0.05). Par contre, le BQ n'a pu discriminer les patients ayant eu une ostéotomie de ceux ayant eu ostéotomie et fusion (p=0.14). Les auteurs concluent à la validité de construction du BQ dans ce groupe de patients. La seule critique que l'on peut trouver est qu'il s'agit d'une étude rétrospective et que le BQ a été utilisé dans une autre pathologie que le SCC.
- Padua (73) dans la publication de la traduction du DASH en italien, a administré le DASH et le BQ aux patients qui avaient un SCC. Leur nombre n'est pas précisé. Les corrélations entre DASH et BQ chez ces patients sont de 0.48 pour le BQ symptômes (p<0.02) et de 0.65 pour le BQ fonction (p<0.02).

V.8.5 Conclusions.

Le BQ est un questionnaire bien construit et facile à utiliser. Il a de bonnes propriétés psychométriques tant en ce qui concerne la fiabilité test retest, la cohérence interne et la validité de construction. La sensibilité au changement dans le SCC est bonne et prouvée par plusieurs études. Dans d'autres pathologies les résultats sont plus contradictoires. Ce questionnaire étant spécifique du SCC, je pense qu'il faut réserver son utilisation à cette pathologie. Il devrait s'imposer comme le « gold standard » pour le SCC (5). Pour d'autres pathologies du poignet le DASH et le PRWE ont montrés leurs qualités. Ce questionnaire s'est maintenant imposé comme la référence pour le SCC et permet de comparer des résultats entre équipes de différents pays avec différentes techniques, d'apprécier le résultat d'une intervention chirurgicale, voire d'une intervention thérapeutique médicale sur le SCC. Malheureusement, ce questionnaire n'a pas encore été traduit en français. Le tableau ci-dessous résume les caractéristiques du BQ.

PROPRIETES	BQ
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Questionnaire d'auto évaluation
4) Interprétation des résultats	A parfaire
5) Spécificité	Syndrome du canal carpien
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	anglais
Coût	gratuit
7) Temps demandé	
Administration	5 à 10 minutes
Entraînement du testeur	non
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	Anglais, italien, espagnol, portugais, suédois
Application dans une autre culture	Oui
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	Oui
Interprétation	+/-
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	Pas besoin
Intra observateur	Pas besoin
Test-retest	0.79 à 0.93 selon les études
Cohérence interne	0.89 à0.91 selon les études
3) Validité	
Contenu	Oui
Construction	Oui en partie
Relation avec AVQ	Oui
Relation avec statut physique	Non
Distinction des groupes	Oui
Critères	Pas étudiée
Sensibilité au changement	Oui pour le SCC ? pour d'autres pathologies non
4) Normes	

Tableau n°24: caractéristiques du BQ.

V.9 Le Alderson-McGall hand function questionnaire (AMHFQ).

Il a été publié en 1999 (3) par une équipe canadienne de l'Ontario. Les auteurs sont partis du constat que le BQ, d'après eux, dans sa partie fonction, n'incluait pas toutes les activités manuelles affectées par le SCC. Il manquait les activités motrices fines (manipuler des pièces de monnaie), les activités de résistance (tourner un robinet), les tâches répétitives de dextérité (clavier d'ordinateur) et les postures statiques (tenir un volant). Le but de leur étude était donc de développer un nouveau questionnaire et d'en étudier les propriétés psychométriques. Il faut noter que Levine lors de la publication du BQ avait bien dit que les activités choisies pour la partie fonction étaient des activités courantes que tous les patients étaient censés faire dans la vie quotidienne.

- **Matériel et méthodes :** 17 sujets ont participé à l'étude. Les tests suivants ont été pratiqués : mesure de la force au Jamar, mesure de la force de la pince, mesure de la discrimination de deux points, mesure de la dextérité avec le kit Valpar Work Sample test standardisé , EVA, EVA fonctionnelle de 0 à 10 et enfin questionnaire AMHFQ. Ce questionnaire a donc été fabriqué par les auteurs, d'après une liste d'AVQ et de loisirs. Les items en rapport avec le SCC ont été sélectionnés par les auteurs. Il n'y a donc pas eu d'avis d'autres experts ni de patients. Le nombre total d'items n'est jamais noté dans la publication, il semble y en avoir entre 50 et 60 ! L'échelle de cotation des items va de 7 (pas de difficulté) à 1 (incapable de faire). Le patient a la possibilité de cocher une case en plus s'il ne fait pas l'activité. Le questionnaire demande 10 à 15 minutes pour être rempli.
- **Résultats :**
  La cohérence interne appréciée par le Cronbach alpha est de 0.97.
  La fiabilité test retest est appréciée par les coefficients de corrélation intra classe. Ils sont tous dans le 95ème percentile sauf un.
  La validité de construction était basée sur l'hypothèse que les scores de l'AMHFQ seraient corrélés aux mesures physiques. Des corrélations positives ont été démontrées entre AMHFQ et EVA douleur, EVA fonction, force au grip et à la pince. Elles sont modérées entre 0.36 et 0.38.

Les auteurs dans la discussion ont bien conscience que ce questionnaire est trop long et qu'il va falloir faire une analyse de facteur pour réduire les items. Ils ne veulent pas perdre des renseignements intéressants pour évaluer la fonction manuelle du patient. Cette réduction des items devrait faire l'objet d'une publication qui n'a pas encore été faite 6 ans après.

En conclusion : pour l'instant ce questionnaire ne peut être retenu. Il est trop long pour une évaluation que l'on voudrait rapide, les items n'ont pas fait l'objet d'un avis d'experts ni de patients et sa validité de construction est encore sujette à caution. L'échantillon testé est très faible (17 patients) et la publication définitive de ce questionnaire n'a pas encore été faite à ce jour. Quand il sera publié, il sera alors intéressant de le comparer au BQ.

V.10 Le Hand Function Scoring system (HFS).

Il a été décrit en 1998 par Watts (94), et est utilisé dans le centre de rééducation de la main de Salisbury en Angleterre. C'est un score « maison ». Il repose sur l'évaluation de 25 AVQ qui

mettent en jeu les mains. Ces activités sont cotées de 1 (facile) à 4 (impossible). Le score va donc de 25 à 100. 85 patients ayant des traumatismes variés de la main ont été évalués à T0 et à la fin du traitement. 64 patients ont pu être évalués. Le score moyen à l'admission était de 77 et à la fin de 46 ($p<0.001$), ce qui traduit une amélioration significative.
C'est la seule donnée que l'on peut tirer de cette étude car le HFS n'a pas été validé, on ne connaît pas sa fiabilité et sa mise au point est sujette à caution. Dans l'état actuel des connaissances sur l'HFS on ne peut recommander son utilisation. Aucune autre publication n'a été faite sur le sujet depuis 1998.

V. 11 Le « patient-focused wrist outcome instrument » (PWOI).

V.11.1 Introduction.

Ce questionnaire émane d'un long travail de réflexion et de mise au point de Madame Bialocerkowski, australienne (13, 14, 15,16). Dans un papier de 2000 (14), une revue de tous les instruments d'évaluation du poignet a été effectuée, en se focalisant sur le contenu et la qualité. Seuls le BQ, le PRWE et le SODA montraient d'après elle, une qualité suffisante. Elle estimait que la plupart des outils d'évaluation du poignet négligeaient de mesurer l'impact de la maladie sur l'individu, et que l'outcome n'était généralement pas exprimé en terme de fonction ou en termes pertinents pour chaque individu. De là, elle a mené une réflexion pour développer un nouvel outil d'évaluation sous forme d'un questionnaire qu'elle a publié en 2003.

V.11.2 Développement du questionnaire.

Dans un premier temps, une étude (13) a été effectuée et publiée en 2002. Le but de cette première étude était d'explorer l'étendue des difficultés fonctionnelles et des mécanismes de compensation rapportés par des patients ayant un problème de poignet, dans le but d'avoir une base de données pour construire un outil d'évaluation. Il s'agit d'une étude descriptive basée sur une interview semi structurée des patients. La structure de l'interview comportait six thèmes : informations démographiques, statut pré traumatique du poignet, informations sur le traumatisme, activités difficiles, mécanismes de compensation et problème principal rencontré par le patient avec son poignet.
Des volontaires ont donc été recrutés dans la région d'Adélaïde en Australie ; ils étaient porteurs d'un problème de poignet, unilatéral. L'interview était enregistrée et la question principale était : « comment votre problème de poignet a-t-il affecté votre capacité à effectuer les AVQ ? » Chaque interview était retranscrite par écrit et le patient devait donner son accord sur cette transcription.
42 patients de 18 à 81 ans ont participé à l'étude. Les diagnostics étaient variés : fractures (50%), lésions des tissus mous (33%), canal carpien (12%) et arthrose (5%).
Une partie des patients ont rapporté des difficultés avec les tâches fonctionnelles avant leur blessure au poignet. Après le traumatisme, un large champ de difficultés, intéressant toutes les activités motrices, est reporté. Ces activités ont été classées. La liste des activités difficiles a été comparée à la liste des activités contenues dans les autres outils d'évaluation. La liste était globalement la même que dans les autres outils, mais le large champ d'activités limitées, jugées importantes par les patients dans cette étude, n'est pas retrouvé dans la plupart des autres instruments.
Les mécanismes de compensation sont utilisés par tous les patients : aide par un tiers, utilisation de l'autre main, utilisation d'une autre partie du corps pour porter et serrer.

Le principal problème perçu par le patient était unique dans 45% des cas et la raideur (47%), la douleur (42%) et la faiblesse (11%) étaient les doléances principales des patients. Il faut noter aussi beaucoup de doléances sur le plan émotionnel ou financier en rapport avec cette pathologie du poignet.
Bialocerkowski estime donc que les questionnaires existants ne sont pas à même de couvrir, l'ensemble des activités limitées rapportées par les patients et donc ne sont pas capables de mesurer pleinement le changement dans la capacité. Cette réflexion est le point de départ pour le développement du PWOI.

La publication de ce questionnaire a eu lieu en 2003 (15, 16). Il se compose de deux parties (15). La première partie est le composant standardisé. La sélection s'est faite à partir des activités qui avaient été rapportées difficiles et importantes dans l'étude de 2002. Ces items ont été classés en deux groupes en fonction du caractère essentiel ou non de l'AVQ pour tous les individus. De ces deux groupes, les items non spécifiques ou utilisant une seule main ont été exclus. Pour répondre aux questions, il a été choisi d'évaluer la capacité réelle du patient. Dans la plupart des autres questionnaires, le patient doit évaluer la capacité estimée qu'il a ou qu'il aurait dans une activité donnée (estimé sur une échelle de 1 à 5 ou 7). Dans ce questionnaire, 4 réponses sont possibles :

- Oui= difficulté pour faire l'activité.
- Non= pas de difficulté à faire l'activité.
- Je n'ai pas essayé= je n'ai pas fait l'activité dans la semaine précédente, mais je pourrais normalement la faire (comme avant mon problème de poignet).
- Pas applicable= le patient ne fait jamais cette activité, par exemple le jardinage s'il habite en appartement.

52 items composent ce composant standardisé. Les deux premiers groupes de questions concernent l'état du poignet avant le traumatisme ceci pour établir une mesure de base du poignet, car nous avons vu que 5% des patients dans l'étude de 2002 signalent un problème de poignet pré existant au traumatisme.
Le deuxième composant du questionnaire est le composant individualisé. Ce composant est utile pour mettre en évidence des informations concernant les AVQ prioritaires, les mécanismes compensatoires, ceci à l'échelon individuel. Il a été construit à partir des résultats de l'article de 2002 et de différents questionnaires publiés tels que le « patient generated index » ou le « patient specific index ». Il se compose de cinq parties :

1. Le patient doit nommer cinq activités difficiles et importantes pour lui, quand elles doivent être faites sans mécanismes de compensation.
2. Il doit ensuite évaluer l'importance de la difficulté de chaque activité sur une échelle de 1 (pas de difficulté) à 5 (je ne peux le faire).
3. Il doit ensuite évaluer l'importance de chaque activité de 1 (pas important) à 5 (essentielle).
4. Il doit ensuite décrire le nombre et le type de mécanismes de compensation utilisés pour chaque activité choisie.
5. Enfin, il doit dire s'il utilise une aide spécifique externe pour chaque activité.

Le questionnaire a ensuite été testé chez des patients pour étudier la fiabilité, la validité et la sensibilité au changement.

V.11.3 Propriétés psychométriques.

Une première étude a été pratiquée chez 50 patients avec des pathologies variées. Il a été étudié la validité de contenu et la clarté des items, l'existence ou non d'un effet plafond/plancher et la fiabilité test retest.

- **Validité de contenu et clarté des items :** tous les patients qui ont participé à l'étude, ont estimé que le questionnaire couvrait l'ensemble de la capacité à effectuer les AVQ, ce qui prouve sa validité de contenu. Des changements minimes ont été effectués dans le questionnaire au début du test.
- **Fréquence des réponses :** toutes les catégories de réponses sont retrouvées dans le composant standardisé. Aucun patient n'a rapporté de difficultés avec toutes les AVQ. Ceci indique que les items et les catégories de réponses, sont appropriés à une grande partie des individus avec pathologies de poignet. Pour le composant individualisé, 49 des 50 patients ont nommés des AVQ difficiles et importantes, qui étaient variées et assez différentes de celles du composant standardisé. Les compensations étaient utilisées dans 85% des cas. Ce composant individualisé parait donc important et utile.
- **Fiabilité test retest :** elle est bonne avec un coefficient Kappa=0.85.

Une autre étude (16), du même auteur, s'est attachée à étudier la validité de construction et la sensibilité au changement. 26 patients avec une fracture du poignet ont composé le groupe testé. Les mesures ont été faites à 8, 12 ,18 et 24 semaines post fracture. A chaque visite, les mesures suivantes étaient pratiquées :

1. Le questionnaire PWOI avec un score du composant standard et un score du composant individuel.
2. Le Glasgow Pain Questionnaire, questionnaire sur la douleur.
3. Mesures des amplitudes articulaires du poignet, force au Jamar et force de la pince, mesure de la pression avec les mono filaments de Semmes-Weinstein.

- **Validité de construction :** les deux hypothèses utilisées pour la mettre en évidence étaient : 1) que la capacité d'un individu d'effectuer des AVQ augmente avec le temps, plus on s'éloigne de la fracture. 2) des corrélations faibles ou modérées sont attendues entre le score PWOI et les mobilités du poignet ou la force du grip.
  Elle a été évaluée par la mesure répétée de l'analyse de la variance. Un p< à 0.05 est significatif. Les scores du PWOI changent (s'améliorent) sur une période de 16 semaines et cela de manière significative. Ceci confirme l'hypothèse n°1.
  L'association avec les variables cliniques est faible à modérée sauf pour la force de la pince (r entre -0,43 et -0.50). La force de la pince prédit entre 18 et 25 % de la capacité à effectuer les AVQ. Par contre, il existe une association assez forte entre la douleur et le PWOI (r entre 0.41 à 0.83). La restriction des AVQ causée par la douleur peut prédire 56 % de la capacité à effectuer ces AVQ. Il existe une association modérée entre les deux composants du questionnaire (coefficient de corrélation de Pearson entre 0.39 et 0.83), mais ils mesurent deux aspects différents.
- **Sensibilité au changement :** a été calculée par l'effet de taille. Les effets de taille des deux composants sont de 1.71 pour le composant individuel et de 1.68 pour le composant standardisé. Ceci confirme l'hypothèse selon laquelle le PWOI peut détecter le changement au cours du temps chez des patients ayant eu une fracture du poignet. Comparés à la plupart des mesures habituelles, les deux composants ont été plus sensibles pour détecter le changement à 24 semaines. Il faut tout de même noter un effet de taille de la flexion du poignet à 1.68 et de la force de la pince à 1.27. Ceci confirme aussi l'hypothèse selon laquelle le PWOI est plus à même de détecter un changement que les mesures classiques.

V.11.4 Conclusions.

Le PWOI est un questionnaire intéressant et original. Il est le fruit d'un long travail et d'une réflexion de plusieurs années. Ces principales qualités sont les suivantes :

1. Les items qui le composent sont dérivés de l'interview de patients ayant des pathologies du poignet.
2. Il mesure la capacité réelle des patients.
3. Il permet une comparaison avec l'état de base
4. Il permet d'identifier les priorités pour le traitement, grâce au composant standardisé. Il permet de cibler le traitement sur les items auxquels le patient a répondu oui ou je n'ai pas essayé.
5. Il permet de mettre en évidence les activités importantes et difficiles pour le patient, grâce au composant individualisé.
6. Il évalue les mécanismes de compensation utilisés par le patient.

La plupart de ces qualités n'étaient pas il est vrai, explorées par les autres questionnaires.
Les qualités psychométriques de base semblent bonnes mais il faut noter que la cohérence interne n'a pas été calculée.
Bien sur, cet outil n'est pas encore irréprochable et de nombreuses propriétés sont encore à préciser. Le calcul du score en particulier n'est pas encore standardisé et semble complexe en particulier pour les réponses « pas applicable » et « je n'ai pas essayé ». Pour le moment les auteurs se sont contentés d'utiliser les réponses oui et non. C'est à mon avis l'obstacle principal actuellement. Des recherches complémentaires sont en cours pour le score et pour intégrer les mécanismes de compensation dans le score. De même, d'autres études sont nécessaires pour vérifier la validité de construction, la fiabilité test retest et surtout la sensibilité au changement. Il sera ensuite intéressant de comparer ce questionnaire au PRWE ou au DASH. Le tableau ci-dessous résume les caractéristiques du PWOI.

PROPRIETES	PWOI
UTILITE CLINIQUE	
1) Type des résultats	Quantitatifs et qualitatifs
2) Type des tâches	
Représentatif des AVQ	Oui
Unilatéral	Oui
Bilatéral	Oui
3) Méthode d'administration	Questionnaire d'auto évaluation
4) Interprétation des résultats	Pas encore au point
5) Spécificité	Pathologies du poignet
6) Disponibilité	
Préfabriqué	Oui
Domaine public	Oui
Langue de publication	anglais
Coût	gratuit
7) Temps demandé	
Administration	10 à 15 minutes
Entraînement du testeur	non
8) Acceptabilité	
Description du but	Oui
Approprié à des adultes	Oui
Langue	Anglais
Application dans une autre culture	A priori
STANDARDISATION	
1) Instructions	
Administration	Oui
Calcul du score	non
Interprétation	non
2) Equipement	Préfabriqué
BUTS	Descriptif et évaluatif
PROPRIETES PSYCHOMETRIQUES	
1) Construction des items	Descriptif
2) Fiabilité	
Inter observateurs	Pas besoin
Intra observateur	Pas besoin
Test-retest	Oui mais dans une seule étude
Cohérence interne	?
3) Validité	
Contenu	Oui
Construction	Oui en partie
Relation avec AVQ	Oui
Relation avec statut physique	Non
Distinction des groupes	?
Critères	Pas étudiée
Sensibilité au changement	Oui dans une étude pour fracture poignet
4) Normes	non

Tableau n°25: caractéristiques du PWOI.

V.12 Conclusions.

Ce chapitre sur les questionnaires d'auto évaluation était long et reflète bien l'importance qu'ils prennent actuellement dans l'évaluation fonctionnelle de la main et du membre supérieur en général. Tous ces questionnaires ont été publiés entre 1993 et 2004, soit sur une période de 10 ans. La qualité moyenne des questionnaires est bonne à très bonne car ils ont pour la plupart été développés selon des critères précis et stricts. Il faut retenir que le questionnaire idéal n'existe pas et n'existera sûrement jamais. Il faut toujours faire un compromis entre les informations qu'on veut qu'il apporte et sa longueur. 15 minutes pour le remplir semblent être un maximum et il faut privilégier des questionnaires avec moins de 30 à 40 items, sous peine de voir la motivation du patient pour le remplir fléchir fâcheusement. Il doit être facile à comprendre et à remplir. Idéalement, il devrait pouvoir être rempli par exemple avant une consultation, dans la salle d'attente. Enfin, il doit avoir des qualités psychométriques les meilleures possibles, pour que l'utilisateur soit sûr d'avoir un outil fiable, validé et sensible au changement.
Trois questionnaires me paraissent remplir ces conditions.
Tout d'abord le DASH qui est celui qui est le mieux étudié et dont la mise au point a été la plus rigoureuse. Son succès à travers le monde en est d'ailleurs la preuve. C'est un outil en constante évolution et qui bénéficie d'un suivi régulier par l'équipe qui l'a mis au point. Il est utilisable pour toute pathologie du membre supérieur ce qui en fait un questionnaire générique du membre supérieur. Actuellement c'est le seul à pouvoir avoir ce titre. On peut l'utiliser seul ou en association avec un questionnaire plus spécifique d'une articulation, ou bien avec un questionnaire de santé générale comme le SF 36. Il existe en version française validée ce qui est un avantage supplémentaire. Je ne peux que recommander son utilisation systématique dans une évaluation fonctionnelle de la main en complément bien sûr de l'examen clinique et de tests classiques de la fonction de la main.
Ensuite le PRWE, questionnaire spécifique du poignet, est aussi utilisable de façon sûre et fiable. Il est un peu plus court que le DASH et plus facile à calculer. On peut l'utiliser en association avec le DASH ou bien avec un questionnaire comme le SF36. Il est aussi disponible en français. Il s'adresse surtout à des pathologies orthopédiques.
Enfin, le BQ spécifique du canal carpien, est un outil reconnu, validé, qui est considéré comme le « gold standard » pour cette pathologie. Il n'est malheureusement pas disponible en français.
Trois autres questionnaires me paraissent pouvoir devenir des outils intéressants.
Le MHQ, qui s'intéresse à la main et au poignet, étudie des dimensions différentes du DASH par exemple. Ses propriétés psychométriques demandent à être précisées.
Le POS est conçu pour être administré avant et après un geste chirurgical. Ses propriétés psychométriques sont bonnes mais demandent à être étoffées. Il sera sûrement dans l'avenir un questionnaire intéressant et utile pour la chirurgie.
Le PWOI, propose un concept intéressant pour le poignet, mais ses propriétés demandent à être précisées. C'est un outil intéressant, qui peut aider à guider la thérapeutique médicale, en centrant les interventions thérapeutiques sur les incapacités importantes pour le patient.
Les autres questionnaires ne peuvent être retenus en l'état actuel des données.
J'aimerais faire un plaidoyer pour que ces questionnaires soient utilisés de manière plus large voire systématique par tous les thérapeutes qui s'occupent de la main. Au même titre que l'examen clinique l'utilisation de ces questionnaires devraient devenir une évidence et faire partie intégrante de l'évaluation fonctionnelle de la main.

VI. EVALUATIONS CENTREES SUR L'ACTIVITE PROFESSIONNELLE.

VI.1 Introduction.

Nous avons vu longuement les tests d'évaluation de la fonction de la main et les questionnaires d'auto évaluation. Ces tests sont utiles et le plus souvent suffisants pour des patients sédentaires, retraités ou ayant un travail sans grande demande physique. Par contre, pour une grande partie des patients que nous traitons, qui sont des travailleurs manuels ou de force, l'évaluation doit aller plus loin et coller le plus possible à l'activité professionnelle exercée. Les tests étudiés plus haut se révèlent alors insuffisants, car même si le patient a des capacités retrouvées dans sa vie quotidienne, il n'en va pas forcément de même pour son activité professionnelle. Il faudra alors évaluer le patient en situation de travail, simulée ou réelle. L'évaluation en situation réelle est bien sûr l'idéal, mais ne peut être réalisée facilement dans toutes les structures de réadaptation.
Dans ce chapitre, je vais aborder rapidement dans une première partie les tests standardisés qui évaluent la « capacité fonctionnelle », ces tests sont tous nord américains et nous verrons qu'ils sont peu intéressants car très ciblés sur des travaux répétitifs en usine. Ensuite je présenterai notre expérience à la clinique romande de réadaptation SUVA care, en ce qui concerne l'évaluation fonctionnelle du membre supérieur selon Ishernagen. Enfin je montrerai l'utilité d'une observation dans nos ateliers professionnels intégrés à la clinique et ce que cette observation peut apporter.

VI.2 Tests standardisés centrés sur le membre supérieur (9).

Ils ont été développés aux USA pour essayer de répondre à la question : le patient peut il reprendre son travail antérieur et quand peut il le faire ? Le concept d'évaluation de la capacité fonctionnelle (ECF) a été conçu comme une batterie de tests pour évaluer les capacités fonctionnelles et physiques du patient. L'ECF comprends deux parties : un statut clinique et des tests standardisés et observationnels pour analyser les capacités du patient à faire son travail et à tolérer les contraintes physiques de son travail.
La corrélation de l'ECF aux contraintes imposées par un travail donné est fait via le « dictionary of occupational titles=DOT » qui est une base de données sur l'analyse des différents métiers et leurs contraintes. Le thérapeute peut alors sélectionner les tests qui simulent le mieux les contraintes physiques du métier. Chaque ECF est alors adaptée au patient et à son métier. Il ne devrait donc pas y avoir deux ECF similaires. Outre les capacités manuelles, l'ECF permet par l'observation d'évaluer la position assise, la position debout, la marche, le port de charges et la manutention.
De nombreux tests standardisés existent et il s'agit d'un marché commercial important car ils sont tous soumis au copyright. Nous allons rapidement les passer en revue.

- **Le « VALPAR work samples »:** il est subdivisé en plusieurs tests spécifiques qui sont adaptés aux patients traumatisés de la main. Le calcul du score met l'accent sur la qualité de la tâche et du mouvement et sur le temps. Il existe des normes qui sont valables pour chaque travail. La fiabilité test retest est bonne si la procédure est bien suivie.
    1. **« The Whole body range of motion work sample »:** Il teste les mouvements grossiers du corps combinés avec la dextérité fine bilatérale. Le patient doit mettre et défaire des boulons et écrous qui maintiennent trois formes en plexiglas d'un large support ajustable en hauteur. Le but est de transférer les 3 formes : de la hauteur des épaules au dessus de la tête, de au dessus de la tête jusqu'à la taille, de la ceinture au niveau des genoux et des genoux jusqu'à

hauteur des épaules. Le temps est mesuré et le thérapeute observe le patient (compensations, ergonomie du geste…).

2. **« The Valpar upper extremity range of motion work sample » et « The Valpar small tools » :** ils testent la dextérité bilatérale fine. Le premier test nécessite la manipulation de boulons sur des chevilles à l'intérieur d'une boite. Le deuxième, est basé sur le même concept, mais nécessite des outils pour insérer des vis, des boulons et d'autres éléments. Pour ces deux tests, une variété de postures de poignet est nécessaire et on peut faire aussi le test sans contrôle de la vue.

- **Le « Work evaluation system (WEST) » :** les sous tests WEST 2 et 4 sont utilisables pour les mains traumatisées. Le patient doit travailler contre résistance et ces tests mesurent la capacité du membre supérieur en conjonction avec les mobilités du corps entier. Le WEST 2 a une résistance réglable entre 5 et 30 livres et il peut aussi être utilisé pour faire de l'assemblage/désassemblage. Le WEST 4 mesure la force du membre supérieur et son endurance, dans des tâches contre résistance qui nécessitent de la prono supination de l'avant bras ou des rotations de l'épaule. La procédure est standardisée mais la fiabilité n'est pas connue et il n'y a pas de normes disponibles.
- **Le « BTE (Baltimore therapeutic equipment) Work simulator » :** a été développé à Baltimore. Il simule les mouvements du membre supérieur utiles pour réaliser des tâches au travail. On utilise des outils qui vont reproduire les mouvements spécifiques du membre supérieur durant une activité donnée, et ceci contre une résistance mesurable pendant une durée de temps connue. Avec un logiciel, des calculs compliqués sont pratiqués pour déterminer des coefficients de variations pour certains outils sur le BTE work simulator par rapport à l'effort maximal d'un sujet normal. Le BTE peut donc être utilisé pour évaluer la capacité du patient à effectuer certaines tâches physiques du poste de travail. Un protocole standardisé et des normes existent. La fiabilité a été démontrée, elle est de 0.979 pour la force du grip et de 0.913 pour la force en flexion du poignet.
- **Le « ERGOS WORK SIMULATOR » :** il s'agit d'un test datant de 1993, dont une publication récente (17) a étudié la fiabilité. C'est une marque déposée. Il comprend cinq modules, les modules 2 et 5 concernent le membre supérieur. Le module 2 étudie différents types de mouvements actifs dans différentes positions du corps. Le module 5 étudie des travaux sédentaires. Cette étude n'a étudié que la fiabilité du module 5. le module 5 comprend 7 tests : sensibilité digitale, test du clavier, utilisation des doigts et de la main, force du grip, force de la pince, force des flexions du poignet et force de la prono supination. Pour ce module 5, il existe une fiabilité test retest sur une période 4 semaines chez des sujets sains.

Tous ces tests sont assez proches et reposent sur le même principe. Ils peuvent être utiles dans certaines situations de travail bien précises, par exemple pour quelqu'un qui fait un travail très répétitif en usine avec 3 ou 4 tâches seulement. A ce moment là, il est certainement facile de simuler la ou les tâches grâce à ces différents outils et juger de l'aptitude de l'individu à reprendre son activité. Ces outils pourraient même servir alors à un réentraînement au travail. Mais pour évaluer un individu qui travaille sur un chantier tel un maçon, un charpentier, un menuisier, un peintre, un soudeur, pour prendre les professions du bâtiment les plus fréquentes, je pense que ces outils ne sont pas très utiles et n'apportent pas des renseignements déterminants pour juger de l'aptitude au travail. De plus, ces outils ont des désavantages comme le coût, l'encombrement et leurs propriétés psychométriques sont peu étudiées en particulier la validité de construction. Comme le dit Bear-Lehman (9) : « les outils

et les techniques sont devenues de plus en plus sophistiqués ; cependant, le degré de validité concurrente et de validité prédictive reste bas ».

Je vais aborder maintenant deux types d'évaluations complémentaires que nous pratiquons à la clinique romande de réadaptation SUVA care, et qui paraissent plus intéressantes et plus proches des activités professionnelles du travailleur manuel ou de force. Ces évaluations complètent l'évaluation classique : clinique, tests de dextérité, auto questionnaire.

VI.3 Evaluation des capacités fonctionnelles (ECF) du membre supérieur selon Isernhagen.

Dans notre établissement, nous utilisons l'ECF selon Susan Iserhagen, qui est une ECF validée, standardisée, reproductible, bien connue et utilisée dans de nombreux pays. Cette ECF a été conçue au départ pour évaluer les lombalgiques chroniques. Elle se déroule normalement sur deux jours et comprend de nombreuses épreuves variées qui font intervenir l'ensemble du corps et des articulations. Cette ECF est une marque déposée et son utilisation nécessite un agrément et une formation assez longue pour ceux qui sont amenés à l'administrer, c'est-à-dire les kinésithérapeutes. Elle est très coûteuse en temps et en énergie car après l'administration, il faut faire l'interprétation qui demande beaucoup de travail.
L'objectif principal de toute ECF est de déterminer le plus objectivement possible et en toute sécurité les aptitudes physiques nécessaires à la réalisation d'un travail productif. C'est donc l'évaluation des aptitudes physiques à accomplir des tâches ayant un rapport direct avec celles du métier considéré.
Le DOT, dont nous avons déjà parlé au chapitre précédent, est utile pour définir les tâches effectuées dans chaque profession. A chaque profession est attribué un niveau d'activité :

- Sédentaire : < 5kg.
- Léger : 5-10 kg.
- Moyen : 10-25 kg.
- Lourd : 25-45 kg.
- Très lourd : > 45 kg.

La méthode d'évaluation privilégie l'approche kinésio-physique qui est l'appréciation objective de l'effort maximal qui peut être produit plutôt que le respect de la douleur annoncée (psycho-physique). Le niveau de l'effort physique accompli est déterminé par une observation « objective » selon des critères standardisés (identification de l'effort maximal et sans risque). L'évaluateur et non le patient décide de l'arrêt de l'effort (la fonction gouverne plutôt que la douleur). L'arrêt n'a lieu que si l'effort maximal sans risque est atteint ou si l'évaluateur considère que les mouvements sont exécutés de manière si peu sûre qu'ils menacent la santé du patient. La sécurité d'exécution est un principe prioritaire (examen physique, analyse du mouvement et de l'effort, cardiofréquencemètre). Les limites physiques sont distinguées de l'auto limitation.

Déroulement d'une ECF :

L'ECF du membre supérieur est beaucoup plus courte, puisqu'elle se fait sur 1h30 environ pour le test lui-même, plus 1h30 pour rédiger le rapport.
Le patient commence par remplir le questionnaire PACT (Performance Assessment and Capacity Testing) qui est une auto évaluation de ses propres capacités fonctionnelles. Ce questionnaire a été créé par Leonard et Mary Matheson. Il contient 50 illustrations d'activités de la vie quotidienne ou professionnelle. Il évalue semi quantitativement le handicap fonctionnel perçu par le sujet. Le score obtenu au PACT permet de classer le patient dans les

cinq niveaux d'activités définis plus haut et ensuite de comparer aux véritables capacités démontrées lors du test.
Ensuite un rapide examen clinique dirigé sur les membres supérieurs est pratiqué.
Puis les tests ci-dessous sont administrés.
Les tests retenus pour l'ECF membre supérieur sont :

1. **Lever du sol à hauteur de la taille**. Ce test évalue la capacité de flexion/extension des membres supérieurs (mains, poignets, coudes, épaules et nuque), la préhension et l'inclinaison radio-ulnaire.
2. **Lever de la taille à hauteur de la tête.** Cette activité impose une compression sur les articulations des extrémités supérieures et teste la tolérance à la position des poignets et des doigts. Elle permet également d'évaluer la technique ergonomique de travail adoptée par le sujet et les stratégies adoptées face à ses limitations fonctionnelles. Elle teste aussi la préhension et l'inclinaison radio-ulnaire.
3. **Porter de la main droite/gauche (15 m).** Permet la comparaison avec le côté sain.
4. **Porter devant à deux mains (15m).** Activité la plus utilisée permettant d'évaluer la capacité de la musculature des membres supérieurs dans une activité de type principalement statique.
5. **Travailler au dessus du niveau de la tête (5 minutes).** Activité sans charge permettant d'évaluer l'endurance de la musculature des membres supérieurs et la capacité de coordination.
6. **Pousser/tirer en statique (dynamomètre).** Teste la force maximale ou la capacité de résistance des poignets aussi bien à la compression (pousser) qu'à la distraction (tirer). En règle générale le « tirer » est supérieur au « pousser » mais si c'est le contraire, il faut évaluer les restrictions physiques engendrées par le problème local.
7. **Pousser/tirer en dynamique (9m).** Rajoute la composante de coordination des différents groupes musculaires des membres supérieurs entrant en jeu lors d'une activité maintenue sur une distance.
8. **Marche à 4 pattes (9m).** Evalue la tolérance à la compression des poignets, la capacité à stabiliser les articulations périphériques des membres supérieurs et la position adoptée par les poignets/mains (comparaison).
9. **Monter sur une échelle.** Permet d'impliquer directement la prise de la main dans une activité où la sécurité est de mise. Le test peut également être réalisé avec une charge portée dans la main (saine ou atteinte).

Il faut noter que pour le port de charges, le sujet ne connaît pas le poids des charges qu'il soulève. Des sacs de couleurs différentes sont utilisés (bleu et rouge), qui peuvent contenir des charges variables de 2,5 ou 5 kg..
A la fin du test, une courte synthèse avec le patient est pratiquée et les résultats sont discutés avec lui.
Ensuite, le rapport est rédigé par le kinésithérapeute. On peut effectuer un « job match » qui est un tableau comparatif des performances accomplies au cours de l'évaluation avec les exigences physiques des principales tâches du poste de travail. La description du poste de travail est donnée soit par le sujet, soit par l'employeur, soit par des banques de données officielles répertoriant les principaux postes de travail et leurs caractéristiques.
L'évaluateur après ce test doit faire une révision de cohérence, c'est-à-dire voir s'il existe des incohérences dans les tests. Les critères de cohérence sont les suivants :

- L'examen clinique doit être cohérent avec les capacités fonctionnelles démontrées.
- Le lever horizontalement doit être supérieur au lever sol-taille.
- Le lever taille-tête doit être corrélé avec l'activité prolongée au dessus de la tête.
- La force du tirer doit être supérieure à celle du pousser.

- La fréquence cardiaque doit augmenter en relation avec le poids des charges.
- Le niveau d'effort correspondant au score du questionnaire PACT doit correspondre au niveau d'effort en rapport avec les capacités/performances réalisées pendant les tests.
- On peut comparer la marche à 4 pattes avec le lever taille-tête, pour juger de la mobilité des épaules. Par exemple un sujet dit ne pas pouvoir aller au-delà de 90° d'antéflexion des épaules lors de l'épreuve taille-tête et lors de la marche à 4 pattes la mobilité en antéflexion atteint sans problèmes 120 à 140 °.

Si plus de 2 items sont présents, les résultats sont déclarés incohérents. Cela est discuté avec le patient.
Cette ECF permet aussi d'apprécier la coopération du patient et sa motivation (envie de donner le maximum).

Le rapport de l'ECF va donc comporter :

- Les plaintes principales actuelles.
- Un court résumé du statut locomoteur.
- La notion de coopération ou non, de cohérence ou non.
- Le score du PACT.
- Le niveau général d'effort fourni selon les 5 catégories du DOT.
- Les aptitudes physiques liées au travail du sujet.
- S'il est utile de mettre sur pied une réadaptation ciblée ?
- Propositions d'aménagements simples du poste de travail.

La réalisation d'une ECF demande du temps et de l'entraînement et son interprétation peut être délicate. Nous la réservons à des patients dont les performances aux tests fonctionnels habituels sont cohérentes et dont on pense qu'ils vont pouvoir reprendre une activité professionnelle. Chez ces patients là, elle nous est très utile pour pouvoir se prononcer sur les aptitudes fonctionnelles dans un métier donné, pour fixer des limitations par exemple de port de charges, ou simplement pour redonner confiance au patient en lui montrant où il se situe à un moment donné et lui montrer qu'un progression est possible. L'ECF met en avant les aptitudes plutôt que les incapacités ce qui est plus positif pour le patient et son employeur. Bien sur, cette ECF a des limitations comme toute évaluation (collaboration du patient), mais encore une fois chez un patient collaborant et avec des performances cohérentes, elle apporte des informations utiles. Elle est une des pièces du puzzle qui permet dans les cas difficiles de se prononcer sur l'aptitude à un travail donné.
Nous avons la chance de disposer à la clinique romande de réadaptation SUVA care, d'ateliers professionnels qui permettent une évaluation grandeur nature dans la plupart des professions.

VI.4 L'évaluation aux ateliers professionnels.

Au sein du site de la clinique, nous disposons d'ateliers professionnels avec des maîtres socioprofessionnels, ce qui permet une évaluation dans les principaux métiers surtout du bâtiment. Il y a un atelier bois pour la menuiserie et la charpente ; un atelier fer qui permet de faire des constructions métalliques, de la soudure ; un atelier maçonnerie ; de nombreuses machines pour faire du tournage, fraisage, etc. De nombreux autres métiers peuvent être évalués : peintre, plâtrier, monteur de faux plafonds, jardinier, nettoyeur, etc. Il existe aussi un atelier informatique qui permet de s'initier aux techniques de base et à des logiciels plus complexes tels que le dessin assisté à l'ordinateur (DAO). Nous avons donc la possibilité de

mettre le patient dans une configuration de travail proche de son activité habituelle. Le patient tout en suivant des thérapies classiques peut commencer à retrouver peu à peu un milieu professionnel.
Les ateliers professionnels sont utilisés avec plusieurs buts :

- **Aider à évaluer la capacité résiduelle d'un travailleur dans sa profession**. Après un accident grave de la main chez un maçon par exemple, il est parfois difficile de se prononcer d'emblée sur une reprise dans le travail habituel ou sur l'opportunité d'un reclassement professionnel. Il est alors intéressant de confronter le patient progressivement à son métier et de voir par l'observation directe dans une situation donnée, les limitations et les difficultés du patient. Si cette observation montre des limitations fonctionnelles significatives, il faut alors orienter le patient vers un changement d'activité professionnelle.
- **Réentraînement au travail et remise en confiance**. Ceci est valable pour un patient dont on pense qu'il va pouvoir reprendre son activité professionnelle, mais qui est en arrêt depuis longtemps. Il va alors, en parallèle avec les thérapies, débuter un réentraînement au travail progressif, qui va lui permettre de se situer par rapport à son travail et de reprendre confiance en lui. Ceci doit déboucher ensuite sur une reprise à temps partiel ou à but thérapeutique. L'intérêt de ce réentraînement, c'est qu'il peut être ciblé sur des activités précises, que le patient a des difficultés à effectuer, et qui ont été mises en évidence lors des évaluations par les ergothérapeutes, les physiothérapeutes et l'ECF si elle a été pratiquée (par exemple le travail en hauteur).
- **Evaluer les aptitudes pour une autre profession**. Dans un premier temps, il faut déterminer les aptitudes résiduelles liées au travail, puis ensuite chercher les postes de travail aux exigences correspondantes. C'est le départ d'un reclassement professionnel, où il faut fixer ce que le patient est capable de faire avec ses capacités actuelles, avec quel rendement et quelles limitations ? Le patient peut être testé dans des activités variées, y compris en situation de travail vraies à l'extérieur de la clinique dans des entreprises partenaires, pour des activités qui n'existent pas à la clinique. Si certaines conditions sont remplies, le patient peut même débuter à la clinique sa première phase de réadaptation professionnelle, première étape du reclassement.

Ces ateliers professionnels sont très importants pour nous thérapeutes, car ils sont un maillon incontournable dans le processus de réadaptation des patients en âge de travailler, patients qui représentent la majeure partie des blessés de la main et du poignet. Le but ultime de cette réadaptation est tout de même la reprise d'une activité professionnelle, dans la même profession ou dans une autre adaptée au handicap s'il y en a un. L'évaluation aux ateliers fait partie intégrante d'une bonne évaluation fonctionnelle de la main, au même titre que celles des ergothérapeutes et physiothérapeutes.
Bien sûr, les ateliers professionnels ont leurs limitations comme tout outil d'évaluation. La principale limitation est le patient. S'il n'est pas motivé, que ses performances ne sont pas cohérentes, les ateliers ne serviront qu'à mettre en évidence les discordances et le manque de motivation. Je veux encore insister sur l'intérêt des échanges inter disciplinaires dans cette évaluation globale du patient entre les médecins, les ergothérapeutes, les physiothérapeutes et les ateliers professionnels. Ces échanges ont lieu soit lors du colloque hebdomadaire, soit de façon moins officielle sur le terrain par exemple entre physiothérapeutes et ateliers professionnels. Ils permettent de mettre en évidence les limitations objectives, les discordances et performances incohérentes éventuelles, et de fixer des objectifs de réadaptation adaptés à chaque patient.
La deuxième limitation est que le poste de travail exact du patient ne peut être reproduit à 100% dans nos ateliers. Nous essayons en faisant un profil du poste de travail du patient, de

définir les activités qui peuvent poser problème, en fonction des incapacités actuelles, et d'entraîner spécifiquement le patient dans ces activités problématiques.

En conclusion, je dirais que chez un patient volontaire, motivé et dont les performances sont cohérentes, ces ateliers professionnels sont un formidable outil d'évaluation de la fonction manuelle et de thérapie. Bien sûr, peu de patients ont besoin de ces ateliers sur la quantité de blessés de la main, entre 3 à 5 %, mais ce sont ces patients que nous traitons à la clinique et qui posent le plus de difficultés pour les thérapeutes (médecin traitant, chirurgien de la main).

VII. CONCLUSIONS ET PERSPECTIVES.

Dans ce mémoire, j'ai voulu faire une revue la plus exhaustive possible sur les outils à disposition pour l'évaluation fonctionnelle de la main. J'ai volontairement mis de côté l'examen clinique, les tests de force et de sensibilité, pour m'intéresser aux tests de dextérité, aux auto-questionnaires et à l'évaluation centrée sur le travail. C'est en effet dans ce domaine que depuis 20 ans les progrès les plus importants ont été effectués.
L'évaluation et l'interprétation des résultats sont **complexes** (9,68) car elles dépendent de beaucoup de variables. On peut citer : l'intégrité anatomique, la mobilité, la force musculaire, la sensibilité qui sont des variables cliniques, explorées par un examen clinique soigneux, préalable à toute évaluation fonctionnelle. Il faut rappeler ici que la fonction n'est pas corrélée à l'anatomie (5). D'autres variables entrent en jeu : la coordination, l'âge, le sexe, le statut mental, les comorbidités, le niveau socio culturel, la satisfaction du patient vis-à-vis de son traitement. La douleur est aussi une composante essentielle qui influe sur cette évaluation ; il faudra apprécier la réponse à la douleur du patient et sa peur éventuelle de la douleur. Enfin, une variable fondamentale est la participation et la motivation du patient. Une évaluation fiable n'est pas possible sans cela, puisqu'elle est centrée sur le patient et qu'elle vise à fixer des objectifs thérapeutiques pour essayer d'améliorer sa fonction manuelle. Heureusement, dans la plupart des cas, le patient est partie prenante de son évaluation et de son traitement. Un faible pourcentage de patient va poser des problèmes d'évaluation et de traitement, et c'est pour ceux là que le modèle bio psycho social d'Engel (86) peut nous aider à mieux évaluer et comprendre des comportements qui nous déroutent en pratique quotidienne.
L'évaluation fonctionnelle de la main est **inter disciplinaire**, surtout dans les cas compliqués. Elle fait appel au médecin (chirurgien de la main, rééducateur), à l'ergothérapeute, au kinésithérapeute, en premier lieu. D'autres intervenants peuvent être utiles : psychiatre, psychologue, ateliers professionnels, assistante sociale. Il est utile ici de rappeler les 5 composantes nécessaires et suffisantes d'une réadaptation compréhensive (86):

1. Etablissement d'un plan de traitement centré sur le patient, formulé par le patient et l'équipe de réhabilitation.
2. Les objectifs sont posés et agencés par priorité tout au long du processus inter disciplinaire.
3. La participation du patient est requise pour atteindre les objectifs.
4. Les résultats doivent consister en une amélioration du potentiel personnel du patient.
5. Les mesures d'outcome doivent démontrer une réduction des incapacités, du handicap.

Ces composantes sont valables pour toute réhabilitation, et l'évaluation fonctionnelle en est la première étape. Une évaluation fonctionnelle de la main doit se faire sur des données objectives et des résultats validés et fiables (9). Il faut aussi intégrer des données subjectives, peu ou non quantifiables, qui ont aussi une importance, mais un thérapeute qui utilise des outils valides et fiables augmente la probabilité de prendre des décisions cliniques correctes (9). Pour Bear Lehman (9), la thérapie de la main est à la fois un art et une science. Kimmerle (51) en 2003 a développé le concept de répertoire fonctionnel de la main, dans le but de faire un profil complet des aptitudes manuelles du patient. Le propos est identique à ce que j'ai déjà exposé mais présenté d'une autre manière et ce répertoire s'il paraît intéressant reste à valider
Il faut donc des outils qui permettent de faire une évaluation correcte de base, mais aussi qui puissent servir à mesurer et quantifier les progrès effectués par le patient après une thérapie qu'elle soit médicale ou chirurgicale (mesure d'outcome). J'ai passé en revue ces outils tout

au long du mémoire, en essayant pour chacun d'entre eux de mettre en évidence les qualités et les défauts d'après des données scientifiques. On peut dire que la qualité de ces outils s'est bien améliorée, surtout depuis 15 ans, avec l'application d'une rigueur scientifique dans la mise au point, le développement et le suivi de ces instruments.
Quels sont ceux qui peuvent être retenus ? Pour les tests classiques de dextérité, le Purdue et le Minnesota sont à mon avis nécessaires et suffisants. Pour les tests de dextérité de la vie quotidienne, le bilan 400 points a montré des qualités indéniables, mais aussi des défauts ; à mon avis, les qualités l'emportent sur les défauts. De plus, il est disponible en français. Sa sensibilité au changement reste à démontrer ce qui fait qu'on ne peut encore l'utiliser de façon sûre comme outil d'outcome. Deux autres outils le TEMPA et le SODA ont démontré de bonnes qualités globales, mais il s'agit d'outils spécifiques des personnes âgées pour le premier et de la PR pour le second, qu'il faudrait valider dans d'autres populations. Tous les autres outils n'ont pas montré suffisamment de qualités objectives psychométriques pour pouvoir être utilisés sans doute. Il faut d'ailleurs noter que nombre d'entre eux sont utilisés comme référence (Jebsen test, Sollerman test, ARA) alors même qu'ils n'ont pas les qualités initiales requises.
C'est dans les questionnaires d'auto évaluation que l'on trouve le plus d'outils validés et fiables. Ils ont tous été développés depuis les années 1990. J'ai longuement étudié ces questionnaires et on peut retenir le DASH, le PRWE et le BQ (canal carpien). Trois autres questionnaires ont besoin de données complémentaires pour pouvoir être utilisés sans problèmes (POS, MHQ et le PWOI). Je pense vraiment que ces questionnaires doivent avoir une place de choix dans l'évaluation de la main et du poignet, au même titre que l'examen clinique. Ils devraient être intégrés systématiquement à l'évaluation de base et permettent par leur sensibilité au changement de mesurer l'outcome. Le DASH et le PRWE sont disponibles en français dans des versions validées et fiables. Bien sûr, ils ne sont qu'un élément du puzzle « évaluation » mais ils apportent des renseignements importants et surtout permettent de connaître le point de vue du patient sur son état à un instant t.
J'ai montré que l'évaluation fonctionnelle de la main centrée sur le travail se fait au mieux au sein d'ateliers professionnels, en situation réelle. Cette évaluation n'est pas validée scientifiquement mais avec un patient motivé, cohérent, elle donne dans notre expérience des résultats tout à fait satisfaisants. L'ECF qui est une procédure validée donne aussi de bons renseignements, centrés plutôt sur la fonction globale du membre supérieur. Mais ces procédures d'évaluation (ateliers, ECF) sont réservées à des centres spécialisés.

En conclusion, en 2005, je pense que l'évaluation fonctionnelle de la main et du poignet doit comporter un examen clinique bien conduit, des tests de sensibilité si besoin, une étude de la dextérité par des tests classiques (Purdue et Minnesota), une étude de la dextérité centrée sur les AVQ (Bilan 400 points), un ou deux auto questionnaires spécifiques (DASH et /ou PRWE) et éventuellement un auto-questionnaire de qualité de vie tel que le SF36. La mesure de la force est intégrée dans le bilan 400 points. Je pense qu'avec ce bilan de base une bonne photographie de la fonction manuelle peut être prise, photographie qui doit être réitérée dans un délai variable pour juger de l'efficacité d'une thérapeutique. Cette évaluation ne peut être appliquée à tous les patients, surtout en pratique privée, en raison du temps qu'elle nécessite. L'évaluation parfaite n'existe pas et n'existera sans doute jamais, elle sera toujours une estimation de la fonction manuelle avec ce que ce terme comporte d'imprécisions, mais l'utilisation d'outils validés permettra de réduire cette imprécision.

Je terminerais en évoquant les perspectives dans le domaine de l'évaluation fonctionnelle de la main. Elles sont nombreuses car nous ne sommes qu'au début de l'ère de la médecine fondée sur les preuves.
Tout d'abord, il va falloir perfectionner les outils validés existants, en les étudiant chez différents groupes de patients en fonction des pathologies et des traitements, pour mesurer la sensibilité au changement qui est un paramètre important de l'utilité d'un outil.
Ensuite, d'autres outils sont à créer surtout dans le domaine de la participation du patient et de sa qualité de vie (8,86). La nouvelle classification ICF (International Classification of Functioning, disability and health) de 2001 de l'OMS repose sur un concept qui ouvrira de nouvelles voies à l'évaluation (8, 51). Elle constitue une évolution de la classification ICIDH qui reconnaissait 3 domaines : lésionnel (au niveau de l'organe), fonctionnel (au niveau de la personne) avec l'incapacité et enfin le situationnel (au niveau de la société) avec le handicap. La nouvelle classification reconnaît 4 domaines : fonctions organiques, structures anatomiques, activités et participation, facteurs environnementaux. Elle semble plus complexe mais elle sera la référence en réadaptation dans les années à venir.
Enfin, je pense qu'il faudra développer le domaine de l'évaluation fonctionnelle de la main et du membre supérieur en général dans le contexte professionnel, en développant des évaluations ciblées avec par exemple des tâches définies correspondant à un poste de travail défini. Ceci se fera au mieux dans des ateliers professionnels inclus dans un centre de réadaptation, pour permettre des synergies entre les différents thérapeutes. Le problème sera de valider ces interventions avec des outils qui restent à inventer.

BIBLIOGRAPHIE

1. Aaron D.H, Stegink Jansen C.W. Development of the functional dexterity test (FDT): construction, validity, reliability, and normative data. J Hand Ther 2003; 16(1):12-21.

2. Adams B.D, Grosland N.M, Murphy D.M, McCullough M. Impact of impaired wrist motion on hand and upper-extremity performance. J Hand Surg 2003; 28A (6): 898-903.

3. Alderson M, McGall D. The Alderson-McGall hand function questionnaire for patients with carpal tunnel syndrome: a pilot evaluation of a future outcome measure. J Hand Ther 1999; 12(4):313-322.

4. Amadio P.C.Outcomes assessment in hand surgery. What's new? Clin Plast Surg 1997; 24(1): 191-194.

5. Amadio P.C. Outcome assessment in hand surgery and hand therapy: an update. J Hand Ther 2001; 14 (2): 63-67.

6. Amadio P.C, Silverstein M.D, Ilstrup D.M, Schleck C.D, Jensen L.M. Outcome after Colles frature: the relative responsiveness of three questionnaires and physical examination measures. J Hand Surg 1996; 21A (5): 781-787.

7. Atroshi I, Gummesson C, Andersson B, Dahlgren E, Johansson A.The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire. Acta Orthop Scand 2000; 71(6): 613-618.

8. Barbier O, Penta M, Thonnard J-L. Outcome evaluation of the hand and wrist according to the international classification of functioning,disability,and health. Hand Clin 2003; 19(3): 371-378.

9. Bear-Lehman J, Abreu B.C. Evaluating the hand: issues in reliability and validity. Phys Ther 1989; 69(12): 1025-1033.

10. Beaton D.E, Davis A.M, Hudak P, McConnel S. The DASH (disabilities of the arm, shoulder and hand) outcome measure: what do we know about it now? Br J Hand Ther 2001a;6(4): 109-118.

11. Beaton D.E, Katz J.N, Fossel A.H, Wright J.G, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the disabilities of the arm, shoulder and hand outcome measure in different regions of the upper extremity. J Hand Ther 2001; 14(2):128-146.

12. Bell E, Jurek K, Wilson T. Hand skill. A gauge for treatment. Am J Occup Ther 1976; 30(2):80-86.

13. Bialocerkowski A.E. Difficulties associated with wrist disorders-a qualitative study. Clin Rehabil 2002; 16(4): 429-40.

14. Bialocerkowski A.E, Grimmer K.A, Bain G.I. A systematic review of the content and quality of wrist outcome instruments. Int J Qual Health Care 2000; 12(2): 149-157.

15. Bialocerkowski A.E, Grimmer K.A, Bain G.I. Development of a patient-focused wrist outcome instrument. Hand Clin 2003; 19(3):437-448.

16. Bialocerkowski A.E, Grimmer K.A, Bain G.I. Validity of the patient –focused wrist outcome instrument: do impairments represent functional ability? Hand Clin 2003; 19(3): 449-455.

17. Boadella J.M, Sluiter J.K, Frings-Dresen M.H.W. Reliability of upper extremity tests measured by the Ergos TM work simulator: a pilot study. J Occup Rehab 2003; 13(4): 219-232.

18. Cano S.J, Browne J.P, Lamping D.L, Roberts A.H .N, Mcgrouther D.A, Black N.A. The patient outcomes of surgery-hand/arm (POS-hand/arm): a new patient-based outcome measure. J Hand Surg 2004; 29B (5): 477-485.

19. Carroll D. A quantitative test of upper extremity function. J Chron Dis 1965; 18:479-491.

20. Case-Smith J. Outcomes in hand rehabilitation using occupational therapy services. Am J Occup Ther 2003; 57(5): 499-506.

21. Chung K.C, Hamill J.B, Walters M.R, Hayward R.A.The Michigan hand outcomes questionnaire (MHQ): assessment of responsiveness to clinical change. Ann Plast Surg 1999; 42(6):619-622.

22. Chung K.C, Pillsbury M.S, Walters M.R, Hayward R.A, Arbor A. Reliability and validity testing of the Michigan hand outcomes questionnaire. J Hand Surg 1998; 23A (4):575-587.

23. Comte P-A, Vuadens P. Evaluation fonctionnelle de la main en réadaptation neurologique. Rev Méd Suisse Romande 2001 ; 121(6) : 449-451.

24. Davis A.M, Beaton D.E, Hudak P, Amadio P, Bombardier C, Cole D, Hawker G, Katz J.N, Makela M, Marx R.G, Punnet L, Wright J.G. Measuring disability of the upper extremity: a rationale supporting the use of a regional outcome measure. J Hand Ther 1999; 12(4):269-274.

25. Delprat J, Ehrler S, Romain M, Xenard J. Bilan de la préhension. Encycl Méd Chir (Editions scientifiques et médicales Elsevier SAS, Paris, Tous droits réservés), Kinésithérapie-Médecine physique-Réadaptation, 26-008-D-20, 2002, 16p.

26. Desrosiers J, Hébert R, Bravo G, Dutil E. Upper extremity performance test for the elderly (TEMPA): normative data and correlates with sensorimotor parameters. Arch Phys Med Rehabil 1995; 76(12):1125-1129.

27. Draviaraj KP, Chakrabarti I. Functional outcome after surgery for Dupuytren's contracture: a prospective study. J Hand Surg (Am) 2004; 29(5): 804-8.

28. Dubert T, Voche P, Dumontier C, Dinh A. Le questionnaire DASH. Adaptation française d'un outil d'évaluation international. Chir main 2001 ; 20(4) : 294-302.

29. Durand M-J, Vachon B, Hong Q.N, Loisel P. The cross-cultural adaptation of the DASH questionnaire in canadian french. J Hand Ther 2005; 18(1):34-39.

30. Fess E.E. The need for reliability and validity in hand assessment instruments. J Hand Surg 1986; 11A (5): 621-623.

31. Gable C, Chau N, Xenard J, Andre J.M. Accords inter-juges d'une mesure fonctionnelle analytique du membre supérieur. Rev épidémiol Santé Publique 1993 ; 41(1) : 72-83.

32. Gable C, Xenard J, Andre J-M. Bilan 400 points. Journal d'ergothérapie 1996 ; 18(2) : 47-58.

33. Gable C, Xenard J, Makiela E, Chau N. Evaluation fonctionnelle de la main. Bilan 400 points et tests chiffrés. Ann Réadaptation Méd Phys 1997 ; 40 :95-101.

34. Gay R.E, Amadio P.C, Johnson J.C. Comparative responsiveness of the disabilities of the arm, shoulder, and hand, the carpal tunnel questionnaire, and the SF-36 to clinical change after carpal tunnel release. J Hand Surg 2003, 28A (2): 250-254.

35. Germann G, Harth A, Wind G, Demir E. Standardisierung und validierung der deutschen version 2.0 des "disability of arm, shoulder , hand" (DASH)-fragebogens zur outcome-messung an der oberen extremität. Unfallchirurg 2003; 106(1):13-19.

36. Germann G, Wind G, Harth A. Der DASH-fragebogen: ein neues instrument zur beurteilung von behandlungsergebnissen an der oberen extremität. Handchir Mikrochir Plast Chir 1999; 31(3): 149-152.

37. Greenseid D.Z, McCormack R.M. Functional hand testing: a profile evaluation. Plast Reconstr Surg 1968; 42(6): 567-571.

38. Greenslade J.R, Mehta R.L, Belward P, Warwick D.J. DASH and Boston questionnaire assessment of carpal tunnel syndrome outcome: what is the responsiveness of an outcome questionnaire? J Hand Surg 2004; 29B (2): 159-164.

39. Gummesson C, Atroshi I, Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. BMC Musculoskelet Disord 2003 Jun 16; 4(1): 11.

40. Hackel M.E, Wolfe G.A, Bang S.M, Canfield J.S. Changes in hand function in the aging adult as determined by the Jebsen test of hand function. Phys Ther 1992; 72(5):373-377.

41. Hudak P.L, Amadio P.C, Bombardier C and the upper extremity collaborative group (UECG). Development of an upper extremity outcome measure: the DASH (disabilities of the arm shoulder and hand). Am J Ind Med 1996; 29(6): 602-608.

42. Jacobson-Sollerman C, Sperling L. Grip function of the healthy hand in a standardized hand function test. A study of the Rancho Los Amigos test. Scan J Rehabil Med 1977; 9(3):123-129.

43. Jacquot L. Evaluation fonctionnelle de la main. Etude synthétique. Résumé de mémoire en vue de l'obtention du diplôme interuniversitaire en rééducation de la main et appareillage. J Ergother 1998 ; 20(3) : 108-112.

44. Jain R, Hudak P.L, Bowen C.V.A. Validity of health status measures in patients with ulnar wrist disorders. J Hand Ther 2001; 14(2):147-153.

45. Jarus T, Poremba R. Hand function evaluation: a factor study. Am J Occup Ther 1993; 47(5): 439-443.

46. Jebsen R.H, Taylor N, Trieschmann R.B, Trotter M.J, Howard L.A. An objective and standardized test of hand function. Arch Phys Med Rehabil 1969; 50(6):311-319.

47. Jester A, Harth A, Wind G, Germann G. Der "disability of shoulder arms and hand"-fragebogen als instrument zur identification diagnosespezifischer aktivitätsprofile. Unfallchirurg 2003; 106(10): 834-838.

48. Jester A, Harth A, Wind D, Germann G, Sauerbier M. Disabilities of the arm, shoulder and hand (DASH) questionnaire: determining functional activity profiles in patients with upper extremity disorders. J Hand Surg 2005; 30B (1): 23-28.

49. Jones L.A. The assessment of hand function: a critical review of techniques. J Hand Surg 1989; 14A (2 pt 1): 221-228.

50. Karnezis I.A, Fragkiadakis E.G. Association between objective clinical variables and patient-rated disability of the wrist. J Bone Joint Surg 2002; 84B (7): 967-970.

51. Kimmerle M, Mainwaring L, Borenstein M. The functional repertoire of the hand and its application to assessment. Am J Occup Ther 2003; 57(5): 489-498.

52. Kotsis SV, Chung KC. Responsiveness of the Michigan hand outcomes questionnaire and the disabilities of the arm, shoulder and hand questionnaire in carpal tunnel surgery. J Hand Surg (Am) 2005; 30(1): 81-6.

53. Krimmer H, Wiemer P, Kalb K. Vergleichende ergebnisbewertung am handgelenk-mediokarpale teilarthrodese und totalarthodese. Handchir Mikrochir Plast Chir 2000; 32(6): 369-374.

54. Lee E.W.C, Chung M.M.H, Li A.P.S, Lo S.K. Construct validity of the Chinese version of the disabilities of the arm, shoulder and hand questionnaire (DASH-HKPWH). J Hand Surg 2005; 30B: 29-34.

55. Lee E.W.C, Lau J.S.Y, Chung M.M.H, Li A.P.S, and Lo S.K. Evaluation of the Chinese version of the disability of the arm, shoulder and hand (DASH-HKPWH): cross-cultural adaptation process, internal consistency and reliability study. J Hand Ther 2004; 17(4): 417-423.

56. Levine D.W, Simmons B.P, Koris M.J, Daltroy L.H, Hohl G.G, Fossel A.H, Katz J.N. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. J Bone Joint Surg 1993; 75A (11): 1585-1592.

57. Limaye V, Frankham A, Disney A, Pile K. Evaluation of hand function in patients undergoing long term haemodialysis. Ann Rheum Dis 2001; 60(3): 278.80.

58. Luquet C, Chau N, Nadif M, Guillemin F, Gavillot C, Petry D, Moreau T, Bourgkard E, Hecquet B, André J.M, Mur J.M. Unidimensionnalité d'une mesure fonctionnelle pour des malades ayant une lésion du membre supérieur. Rev Epidémiol Santé Publique 1996 ; 44(3) : 248-261.

59. Lyle R.C. A performance test for assessment of upper limb function in physical rehabilitation treatment and research. Int J Rehab Research 1981; 4(4): 483-492.

60. MacDermid J.C. Development of a scale for patient rating of wrist pain and disability. J Hand Ther 1996; 9(2):178-183.

61. MacDermid J.C, Mulè M. Concurrent validity of the NK hand dexterity test. Physiother Res Int 2001; 6(2): 83-93.

62. MacDermid J.C, Richards R.S, Donner A, Bellamy N, Roth J.H. Responsiveness of the short form-36, disability of the arm, shoulder, and hand questionnaire, patient-rated wrist evaluation, and physical impairment measurements in evaluating recovery after a distal radius fracture. J Hand Surg 2000; 25A (2): 330-340.

63. MacDermid J.C, Richards R.S, Roth J.H. Distal radius fracture: a prospective outcome study of 275 patients. J Hand Ther 2001; 14(2): 154-169.

64. MacDermid J.C, Roth J.H, Richards R.S. Pain and disability reported in the year following a distal radius fracture: a cohort study. BMC Musculoskeletal Disorders 2003 Oct 31; 4(1):24.

65. MacDermid J.C, Tottenham V. Responsiveness of the disability of the arm, shoulder, and hand (DASH) and patient-rated wrist/hand evaluation (PRWHE) in evaluating change after hand therapy. J Hand Ther 2004; 17(1):18-23.

66. MacDermid J.C, Turgeon T, Richards R.S, Beadle M, Roth J.H. Patient rating of wrist pain and disability: a reliable and valid measurement tool. J Orthop Trauma 1998; 12(8): 577-586.

67. Macey A.C, Burke F.D. Outcomes of hand surgery. J Hand Surg 1995; 20B (6): 841-855.

68. McPhee S.D. Functional hand evaluations: a review. Am J Occup Ther 1987; 41(3):158-163.

69. Navsarikar A, Gladman D.D, Husted J.A, Cook R.J. Validity assessment of the disabilities of arm, shoulder, and hand questionnaire (DASH) for patients with psoriatic arthritis. J Rheumatol 1999; 26(10): 2191-2194.

70. O'Connor D, Kortman B, Smith A, Ahern M, Smith M, Krishnan J. Correlation between objective and subjective measures of hand function in patients with rheumatoid arthritis. J Hand Ther 1999; 12(4): 323-9.

71. Offenbächer M, Ewert T, Sangha O, Stucki G. Validation of a German version of the disabilities of arm, shoulder, and hand questionnaire (DASH-G). J Rheumatol 2002; 29(2): 401-402.

72. Offenbächer M, Ewert T, Sangha O, Stucki G. Validation of a German version of the "disabilities of arm, shoulder and hand" questionnaire(DASH-G). Z Rheumatol 2003; 62(2): 168-177.

73. Padua R, Padua L, Ceccarelli E, Romanini E, Zanoli G, Amadio P.C, Campi A. Italian version of the disability of the arm, shoulder and hand (DASH) questionnaire. Cross-cultural adaptation and validation. J Hand Surg 2003; 28B (2): 179-186.

74. Penta M, Thonnard J-L, Tesio L. ABILHAND: a Rasch-built measure of manual ability. Arch Phys Med Rehabil 1998; 79(9):1038-1042.

75. Pranski G, Feuerstein M, Himmelstein J, Katz J.N, Vickers-Lathi M. Measuring functional outcomes in work-related upper extremity disorders. JOEM 1997; 39(12): 1195-1202.

76. Rosales R.S, Delgado E.B, Diez de la Lastra-Bosch I. Evaluation of the Spanish version of the DASH and carpal tunnel syndrome health-related quality-of-life instruments : cross-cultural adaptation process and reliability. J Hand Surg 2002 ; 27A (2) :334-343.

77. Rudman D, Hannah S. An instrument evaluation framework: description and application to assessments of hand function. J Hand Ther 1998; 11(4): 266-277.

78. Schuind F.A, Mouraux D, Robert C, Brassinne E, Rémy P, Salvia P, Meyer A, Moulart F, Burny F. Functional and outcome evaluation of the hand and wrist. Hand Clin 2003; 19(3):361-369.

79. Sharma S, Schumacher H.R, McLellan A.T. Evaluation of the Jebsen hand function test for use in patients with rheumatoid arthritis. Arthitis Care Res 1994; 7(1): 16-19.

80. Sinha R, Cresswell TR, Mason R, Chakrabarti I. Functional benefit of Dupuytren's surgery. J Hand Surg (Br) 2002; 27(4): 378-81.

81. Smith H.B. Smith hand evaluation. Am J Occup Ther 1973; 27(5): 244-251.

82. Sollerman C, Ejeskar A. Sollerman hand function test. A standardised method and its use in tetraplegic patients. Scand J Plast Surg Hand Surg 1995; 29(2): 167-76.

83. Solway S, Beaton D.E, McConnell S, Bombardier C. The DASH* outcome measure user's manual. Second edition. Toronto, Ontario: institute for work and health, 2002.

84. SooHoo N.F, McDonald A.P, Seiler III J.G, McGillivary G.R. Evaluation of the construct validity of the DASH questionnaire by correlation to the SF-36. J Hand Surg 2002; 27A (3): 537-541.

85. Stern E.B. Stability of the Jebsen-Taylor hand function test across three test sessions. Am J Occup Ther 1992; 46(7): 647-649.

86. Stiens S.A, Haselkorn J.K, Peters D.J, Goldstein B. Rehabilitation intervention for patients with upper extremity dysfunction: challenges of outcome evaluation. Am J Ind Med 1996; 29(6): 590-601.

87. Turgeon T.R, MacDermid J.C, Roth J.H. Reliability of the NK dexterity board. J Hand Ther 1999; 12(1): 7-15.

88. Van der Lee JH, De Groot V, Beckerman H, Wagenaar RC, Lankhorst GJ, Bouter LM. The intra-and interrater reliability of the action research arm test: a practical test of upper extremity function in patients with stroke. Arch Phys Med Rehabil 2001; 82(1): 14-9.

89. Van der Lee JH, Roorda LD, Beckerman H, Lankhorst GJ, Bouter LM. Improving the action research arm test: a unidimensional hierarchical scale. Clin Rehabil 2002; 16(6): 646-53.

90. Van Lankveld W.G.J.M, Graff M.J.L, van't Pad Bosch P.J.I. The short version of the sequential occupational dexterity assessment based on individual tasks' sensitivity to change. Arthitis Care Res 1999; 12(6):417-424.

91. Van Lankveld W, van't Pad Bosch P, Bakker J, Terwindt S, Franssen M, van Riel P. Sequential occupational dexterity assessment (SODA): a new test to measure hand disability. J Hand Ther 1996; 9(1):27-32.

92. Veehof M.M, Sleegers E.J.A, van veldhoven N.H.M.J, Schuurman A.H, van Meeteren N.L.U. Psychometric qualities of the Dutch language version of the disabilities of the arm, shoulder, and hand questionnaire (DASH-DLV). J Hand Ther 2002; 15(4): 347-354.

93. Voche P, Dubert T, Laffargue C, Gosp-Server A. Auto-évaluation de la douleur et de la fonction du poignet. Proposition d'une version française d'un questionnaire

nord-américain et essai préliminaire. Rev Chir Orthop Reparatrice Appar Mot 2003 ; 89(5) : 443-448.

94. Watts A.M.I, Greenstock M, Cole R.P. Outcome following the rehabilitation of hand trauma patients. The importance of a subjective functional assessment. J Hand Surg 1998; 23B (4): 485-489.

95. Westphal T, Platek S, Schubert S, Schuschke T, Winckler S. Reliabilität und validität des fragebogens "upper-limb-DASH" bei patienten mit distalen radiusfrakturen. Z Orthop Ihre Grenzgeb 2002; 140(4):447-451.

96. Xenard J. Evaluation fonctionnelle de la main. in Monographie du GEM n°26, Réadaptation de la main, Sassoon D et Romain M, Paris expansion scientifique publications 1999, pp73-79.